Crucial steps on the way to a comprehensive structural understanding of the plant proteins EDS1 and PAD4 and their role in innate immunity

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Martin Voß

aus Goch

Köln, Februar 2022

Tag der Disputation: 14.04.2022

Berichterstatter:	Prof. Dr. Karsten Niefind
(Gutachter)	Prof. Dr. Ulrich Baumann

Die in dieser Dissertation dargestellten Arbeiten wurden am Institut für Biochemie (Department für Chemie) der Universität zu Köln (Zülpicherstraße 47, 50674 Köln) unter Aufsicht von Prof. Dr. Karsten Niefind angefertigt.

The work described in this thesis was performed at the institute of biochemistry of the university of Cologne under supervision of Prof. Dr. Karsten Niefind.

Teile der Arbeit wurden bereits veröffentlicht/ Parts of this work have already been published separately.

Voss, Martin; Toelzer, Christine; Bhandari, Deepak D.; Parker, Jane E.; Niefind, Karsten (2019): Arabidopsis immunity regulator EDS1 in a PAD4/SAG101-unbound form is a monomer with an inherently inactive conformation. In: *Journal of structural biology* 208 (3)

Table of contents

Acknowledgements

First and foremost, I want to thank my supervisor Prof. Karsten Niefind for giving me the opportunity to write this thesis. Karsten, I am thankful for your support, scientific discussions, and the freedom to employ my own ideas on how this thesis proceeded. Furthermore, I am thankful for the opportunity to participate in several very fruitful workshops and practical courses, even to topics not directly touched by this thesis.
Secondly, I want to thank Prof. Jane Parker for her scientific input and the opportunity to present and discuss my progress with her research group. Often your input helped to see my results in a biological background, that tends to be forgotten while looking at structures and crystals in a lab that has probably never hosted a real Arabidopsis.
The success of this thesis heavily relied on the support of Prof. Ulrich Baumann. Thank you Uli for your scientific input and for allowing us to take part in your group seminar. A further thank you for graciously allowing us Niefinds to use your equipment and machines. Both Uli and Jane were part of my thesis advisory committee, thank you!

I want to thank Dirk Lindenblatt for his invaluable support during this thesis. Thank you for many fruitful discussions about failed experiments, for listening patiently to me whine about failed gels, bacteria refusing to grow and many a failed purification. Thank you for many nights spend in Grenoble at the, usually faulty, ESRF beamlines. And of course, thank you for producing our awesome element collection (if it ever gets finished).

A special thank you to Dirk for randomly selecting the nanobody 1AT73 that resulted in the best resolution of a EDS1 structure!

Additionally, I want to say thank you to Jennifer Hochscherf for her support and advice.

A big thank you to Christine Tölzer whose introduction to EDS1, nanobodies and the quirks of our lab saved me a lot of time.

I am grateful to the undergrad students I had the pleasure to supervise and whose work helped to create the results presented in this thesis including Alexander Rothemann, Kübra Solak, Attila Szemenyei and Lars Pantenburg. I am grateful to everyone who made working in the Niefind lab a pleasant time including Jenni, Dirk, Christine, Merlin, Christian, Alex, Bianca and many more.

Furthermore, I am grateful for the members of the Baumann lab for scientific discussions, nice chats at the coffee machine and providing a positive and cheerful atmosphere in the institute. Thank you Uli, Trine, Jan, Fabian, Elena, Elena, Doro, Christian, Sabine, Jacky, Matthias, Schirin and Sandra.

Zuletzt möchte ich meiner Familie und meinen Freunden danken, ohne deren Hilfe mein Studium und jetzt meine Promotion nicht möglich gewesen wäre. Vielen Dank an meine Mutter Jutta und meinen Vater Josef für die unschätzbare Unterstützung auf meinem Lebensweg. Das gleiche gilt meinem Bruder Bastian, sowie meinen Freunden, Bekannten und Kegelbrüdern insbesondere, Ati und Gloria.

Abbreviations

Aminoacids

Amino acid	Three letter code	One letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Asp	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophane	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Further abbreviations

Å	angström 1Å=0.1 nm
ADP-ribose	adenosintriphosphate-ribose
ADR1	activated disease resistance 1
AvrRPS4	avirulence RPS4
BAK1	BRI1 associated receptor kinase
BiFc	bimolecular fluorescence completement
BSA	bovine serum albumine
CC	coiled-coil domain
CC-NB-LRR	coiled-coil-nucleotide binding-leucin rich repeat immune receptor
CC-R	coiled-coil RPW8 like domain
cDNA	complementary DNA
CDR	complementary determining region
CNL	CC-NB-LRR protein
Crispr Cas	clustered regulary interspaced short palindromic repeats/ Crispr-associated protein
cryo-EM	cryo-electron microscopy
Da/kDa	dalton/kilodalton 1Da=1g/mol
dATP	desoxyadenosine
DELLA	protein family named after the typical DELLA sequence
DMF	dimethylformamide
DnaJ	chaperone, also called HSP40
DTT	dithiothreitole

dTTP	desoxythymidine
DWARF14	strigolactone esterase D14
ECL	enhanced chemiluminescence lightning
EDR1	enhanced disease resistance 1
EDS1	enhanced disease susceptibility 1
EIJ1	EDS1-interacting J protein 1
EP-domain	protein domain named for EDS1 and PAD4
ETI	effector triggered immunity
FLAG	protein tag
FLS2	flaggelin sensitive 2
GID1	gibberiline insensitive 1
IgG	immunoglobulin G
IP	immunoprecipitation
IPTG	isopropyl-β-D-thiogalactopyranosid
KAI2	karrikine insensitive 2
LB	lysogeny broth medium
LRR	leucin rich repeat
LYK5	lysM containing receptor like kinase 5
LysM	lysin motif domain
MBP	maltose binding protein
MES	2-(N-morpholino)-ethanesulfonic acid
MPD	2-Methyl-2,4,-pentandiol
mRNA	messenger ribonucleicacid
NAD	nicotinamide adenine dinucleotide
NB	nanobody
Ni-NTA	nickel-nitriloacetic acid
NLR	nucleotide-binding leucin rich repeat
NMD	nonsense mediated mRNA decay
NMR	nuclear magnetic resonance
NPR3/NPR4	NPR1-like protein 3/4
NRG1	N-requirement gene 1
PAD4	phytoalexin deficient 4
PAMP	pathogen associated molecular pattern
PBS3	GH3-like defense gene 1
PCR	polymerase chain reaction
PDB	protein database
PopP2	pseudomonas outer protein P2
PRR	pattern recognition receptor
PTI	PAMP-triggered immunity
PVDF	polyvinylidenefluoride
RNA	ribonucleicacid
RPS4	resistance to Pseudomonas syringae 4
RPS6	resistance to Pseudomonas syringae 6
RPW8	resistance to powdery mildew 8
RRS1	resistance to Ralstonia solanacerarum 1
SAG101	senescence-associated gene 101
SAXS	small angle X-ray scattering
SDS	sodium-dodecylsulfate
SDS-PAGE	SDS-polyacrylamide gelelectrophoresis

SMG7	nonsense mediated mRNA decay factor
SNC1	suppresor of NPR1.1
SOBIR1	suppresor of BIR1
SOC	superoptimal broth medium with catabolite repression
SRFR1	suppresor of RPS4-RLD 1
TB	terrific broth
TBST	Tris buffered saline with Tween20
TECEP	Tris(2-carboxyethyl)phosphin
TEMED	tetramethyldiamine
TIR	toll-interleukin receptor 1 like domain
TIR-NB-LRR	TIR nucleotide binding leucin rich repeat immune receptor
TNL	TIR-NB-LRR protein
UV-VIS	ultraviolet and visible light spectrum
v/v	volume/volume
VICTR	variation in compound triggered root growth response
w/v	weight/volume
WRKY	family of transcription factors named after protein sequence
xg	times gravity
Y2H	yeast-two-hybrid
YFP	yellow fluorescence protein
ZAR1	HopZ-activated-resistance 1

Zusammenfassung

Die Infektion von Pflanzen mit Krankheitserregern kann eine ernsthafte Bedrohung für die weltweite Lebensmittelversorgung darstellen. Daher ist die Schaffung resistenter Sorten entweder durch Züchtung oder künstliche genetische Veränderung eine dringende Aufgabe für Forscher und Landwirte weltweit. Um resistente oder widerstandsfähigere Arten gezielt zu entwickeln, ist ein tiefes Verständnis der natürlichen pflanzlichen Immunantwort erforderlich.

Pflanzen verteidigen sich gegen Krankheitserreger mit einem komplizierten, zweischichtigen Abwehrsystem. Die erste Verteidigungslinie besteht aus membranständigen Rezeptoren, die konservierte molekulare Muster von Krankheitserregern wie Flagellin oder Chitin erkennen und die Immunreaktion auslösen. Angepasste Krankheitserreger haben Wege gefunden, die erste Verteidigungslinie zu überwinden, indem sie Effektorproteine in die Zelle injizieren, die die erste Immunantwort blockieren. Als Reaktion darauf haben Pflanzen eine Reihe von löslichen zytoplasmatischen Rezeptoren entwickelt, die die Effektoren entweder direkt oder indirekt erkennen. Die daraus resultierende Immunreaktion führt häufig zum kontrollierten Absterben der infizierten Zelle, um das Wachstum des Pathogens zu begrenzen.

Der Gegenstand dieser Arbeit ist die EDS1-Proteinfamilie mit ihren Mitgliedern „enhanced disease susceptibility 1“ (EDS1), „phytoalexin deficient 4“ (PAD4) und „senescence associated gene 101“ (SAG101). Sie sind zentrale Punkte des pflanzlichen Immunsystems und sind an der Signalübertragung zwischen Rezeptoren und Immunantwort beteiligt. Ein Knockout von EDS1, PAD4 und in geringerem Maße von SAG101 führt zu einer hohen Anfälligkeit der Pflanzen gegenüber einer Vielzahl von Krankheitserregern wie Bakterien, Pilzen und Insekten. Trotz der wesentlichen Rolle von EDS1, PAD4 und SAG101 in der Pflanzenimmunität ist ihre Funktion noch nicht bekannt. EDS1 bildet mit PAD4 oder SAG101 Heterodimere, die unterhalb der Immunrezeptoren beider Abwehrschichten wirken. Alle drei Proteine bestehen aus einer N-terminalen lipaseähnlichen Domäne mit α/β-Hydrolasefaltung und einer C-terminalen einzigartigen Domäne, die nach EDS1 und PAD4 als EP-Domäne bezeichnet wird. Mit Ausnahme einer Struktur des EDS1-SAG101-Dimers aus *A.thaliana* (atEDS1/atSAG101) sind keine strukturellen Informationen über die Proteine bekannt. Frühere Versuche, eine Kristallstruktur von ungebundenem EDS1 zu gewinnen, scheiterten an der geringen Beugungsqualität der gewonnenen Kristalle. In dieser Arbeit habe ich Nanokörper aus *Llama glama* gegen atEDS1 als Kristallisations-Chaperone eingesetzt, um eine Struktur von ungebundenem atEDS1 zu erhalten. Die Kristallisationsversuche waren erfolgreich und ergaben drei Strukturen von atEDS1, die an verschiedene Nanokörper gebunden waren. In Kombination mit zuvor durchgeführten SAXS-Experimenten und *in planta* IP-Daten, die von Kooperationspartnern gesammelt wurden, konnte ich die Quartärstruktur von ungebundenem EDS1 sicher als monomer bestimmen. Frühere Studien, die

über eine Selbstwechselwirkung von atEDS1 berichteten, lassen sich durch große nicht-biologische Kristallkontakte und eine Tendenz des Proteins zur Bildung von Aggregaten erklären.

Da atPAD4 nicht löslich in *E.coli* exprimiert werden kann, habe ich Sequenzen aus verschiedenen Organismen verwendet, um lösliches Protein für die Charakterisierung von PAD4 *in vitro* zu gewinnen. Ein PAD4-Ortholog aus *Vitis vinifera* (vvPAD4) konnte in hohen Mengen in *E.coli* exprimiert werden. Diese Arbeit ist die erste Veröffentlichung, die die Herstellung einer rekombinanten PAD4-Variante beschreibt. Bei dem Versuch, einen EDS1/PAD4-Komplex aus *Vitis vinifera* (vvEDS1/vvPAD4) herzustellen, wurde eine überraschende neue Spleißvariante von vvEDS1 entdeckt, die erfolgreich kristallisiert und deren Struktur gelöst wurde. Diese Entdeckung wirft mehrere neue Fragen über die Regulierung und alternative Spleißvarianten von EDS1 auf

Summary

Infection of cultures of food-crops with pathogens can be a serious threat to the worldwide food supply. Therefore, the creation of resistant varieties, either via breeding or artificial genetic modification, is a pressing task for researchers and farmers worldwide. To rationally design resistant or more resilient species a deep understanding of the natural plant immune response is needed.

Plants defend themselves against pathogens using an intricate two layered defence system. The first line of defence consists of membrane standing receptors that recognize conserved molecular pattern associated with pathogens like flagellin or chitin and trigger the immune response. Adapted pathogens have found ways to overcome the first line of defence by injecting effector proteins in the cell that block the initial immune response. In response, plants developed a series of soluble cytoplasmatic receptors that recognize the effectors, either directly or indirectly. The resulting immune response often leads to the controlled death of the infected cell to limit pathogen growth.

The object of this thesis is the EDS1 protein family with its members enhanced disease susceptibility 1 (EDS1), phytoalexin deficient 4 (PAD4) and senescence associated gene 101 (SAG101). They are central points of the plant immune system and are involved in the signalling between receptors and immune response. A knockout of EDS1, PAD4 and to a lesser degree SAG101 leads to a high susceptibility of the plants against a broad variety of pathogens including bacteria, fungi and feeding insects. Despite the essential role of EDS1, PAD4 and SAG101 in the plant immunity, their function is not yet known. EDS1 forms mutually exclusive heterodimers with PAD4 or SAG101 which act downstream of the immune receptors of both layers of defence. All three protein consist of an N-terminal lipase-like domain with α/β-hydrolase fold and a unique C-terminal domain named after EDS1 and PAD4 as EP-domain. Except a structure of the EDS1-SAG101 dimer from *A.thaliana* (atEDS1/atSAG101), no structural information is known about the proteins. Previous attempts to gain a crystal structure of unbound EDS1 failed due to the low diffraction quality of the crystals gained. In this thesis, I used nanobodies raised from *Llama glama* against atEDS1 as crystallisation chaperones to gain a structure of unbound atEDS1. Crystallisation attempts were successful and yielded three structures of atEDS1 bound to different nanobodies. Combined with previously performed SAXS experiments and *in planta* IP data gathered by cooperation partners, I was able to confidently determine the quaternary structure of unbound EDS1 to be monomeric. Previous studies reporting an atEDS1 self-interactions can be explained by large non-biological crystal contacts and a tendency of the protein to forms aggregates.

As atPAD4 cannot be expressed in soluble form in *E.coli,* I used sequences from different organisms to gain soluble protein for characterisation *in vitro*. An PAD4 orthologue from *Vitis vinifera* (vvPAD4) could be expressed in high amounts in *E.coli*. This thesis is the first report of successfully producing a recombinant PAD4 variant. The attempt to produce a complex of EDS1/PAD4 from *Vitis vinifera* (vvPAD4/vvEDS1) revealed a surprising new splice variant of vvEDS1 that was successfully crystallised and structurally characterized. This finding raises several new questions about the regulation and alternative splice variants of EDS1

1.Introduction

1.1. Social and ecological damage of plant pathogens.

According to data provided by the United Nations, the world is populated by 7.8 billion people as of 2021. The population is expected to grow to 9.7 billion people until the year 2050[1]. At the same time as of 2019, 690 million people (8.9% of the world population at that time) suffer from undernourishment [2]. To feed the drastically increasing world's population that is, even now, suffering in parts from a lack of nourishment, will be one of the large problems of the future. This development will challenge the agriculture worldwide to steadily grow plants of high nutritional value in reliably high yields. Beside the effects of the climate change, the efficient control of pests and pathogens is one of the main challenge farmers faces. Plants as sessile organisms are constantly threatened by a host of pests and pathogens including bacteria, virus, fungi, feeding insects and parasitic plants. The economical and subsequently social loss due to an infected culture can be devastating, for example the parasitic plants of the genus *Striga* (witchweed) is estimated to cause a yearly loss of 7-10 billion US$ and to influence the life of 300 million people (Jamil et al. 2021). Figure 1 shows the global production per person as well as the yield loss in percentage, both as average for the years 2010-2014 for some of the major food crops. The crops were selected by the authors as they provide a major part of the worldwide calorie intake, these are in detail: wheat (18.3%), rice (18.9%), maize (5.4%), potato (2.2%) and soybean (3.3%, Savary et al. 2020). The numbers confer to losses due to pests and pathogens and do not consider other potential reason for reduced yields such as draughts or floods.

These numbers and data demonstrate the need for effective and efficient pest control for plants. One route to ensure continued high yields is the use of pesticides that has been steadily increasing in the last years, as shown in figure 1b. Another avenue is the use of genetically modified plants. The most commonly used modifications convey either resistance to herbicides or resistance against certain pests.

1 https://population.un.org/wpp/

2 http://www.fao.org/3/ca9692en/online/ca9692en.html#

While in Germany no genetically modified plants are grown since 2012[3], the cultivation of such plants is common in many countries. For example, 94% of the soybeans and 92% of the corn grown in the US 2018 were genetically modified[4] .

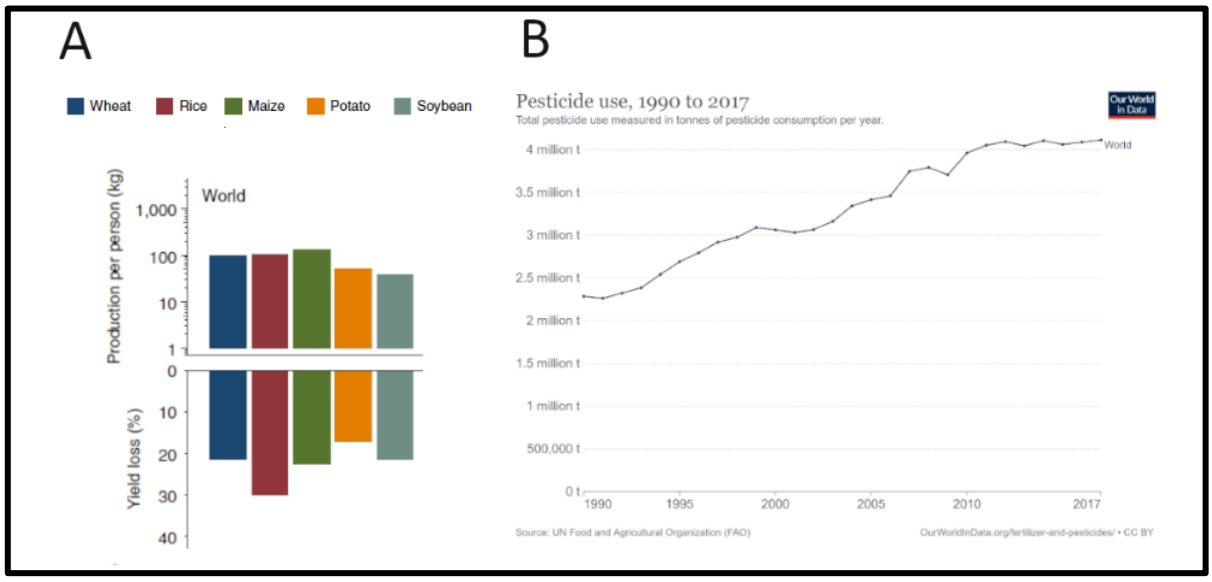

Figure 1: (A) Global production and yield loss due to pathogens in average for the years 2010-2014. The production is shown as crops produced per person in kg. Yield losses are shown in %. Data were collected for five major food crops that make up the majority of the global calorie intake. Modified and taken from Savary et al. 2020 with friendly permission by SpringerNature. (B) Worldwide amount of pesticides used from 1990 to 2017 in million tons.

The targeted manipulation of a plant genome allows for the rapid establishment of a desired phenotype, a distinct advantage to conventional plant breeding techniques. Especially the development of the Crispr-Cas (clustered regularly interspaced short palindromic repeats- Crisps associated protein) method allows the quick and easy manipulation of a plethora of plants. The immense potential of this method, that is not limited to plant breeding, is highlighted by the awarding of the Nobel prize for chemistry to Emmanuelle Charpentier and Jennifer Doudna for the development of the Crispr-Cas method 2020.

3 https://www.bundesregierung.de/breg-de/aktuelles/lebensmittel-in-deutschland-grundsaetzlich-gentechnikfrei-348862

4 https://www.fda.gov/food/agricultural-biotechnology/gmo-crops-animal-food-and-beyond?utm_source=facebook&utm_medium=social&utm_campaign=feedyourmind2020

But while it is, today, possible to manipulate the genome of plants at will, the targeted generation of a distinct phenotype requires detailed knowledge of which gene to target and how a mutation will affect the whole plant. Therefore, to create plants more resistant against certain pathogens it is necessary to know on the one hand, how the plant defends itself against pathogens and on the other hand how a pathogen overcomes this defence.

Beside investigations *in planta*, biochemical methods *in vitro* and especially investigation of key proteins with method of structural biology are necessary to gain an in depth understanding of the plant immune response. The rational mutation and manipulation of plants requires knowledge of how the proteins involved in the response against a pathogen function. Detailed information of structure-function relationships of proteins of interest are an excellent basis to develop rationally designed enhanced plants, that can have a large economic and social impact.

The proteins of the EDS1-family, namely enhanced disease susceptibility 1 (EDS1), phytoalexin deficient 4 (PAD4) and senescence associated gene 101 (SAG101) are central players of the plant immune system. A knockout of the corresponding genes can impair the plant immune response heavily (Dongus & Parker 2021). As such, the investigation of these proteins is essential to decipher the intricacies of plant response against pathogen, and the proteins themselves are potential targets of the rational design of more resistant plants. For example, in wheat, the knock-out of enhanced disease resistance 1 (EDR1), a known antagonist and interaction partner of EDS1 and PAD4 (Neubauer et al. 2020) enhances resistance against *Blumeria graminis,* the cause of powdery mildew (Zhang et al. 2017). Furthermore, it has been shown that the regulation of EDS1 influences the resistance of *Vitis vinifera* against powdery mildew (Gao et al. 2010), thus showing another potential application for genetically modified plants with mutations involving the EDS1 protein family.

1.2. Principles of the plant immune response

The most glaring difference between the immune system of mammals and those of plants is the absence of dedicated immune cells and subsequently an adaptive immune response. While mammals possess specialized cells that e.g., produce antibodies, each plant cell reacts individually in contact with pathogens, although signal molecules can prime other cells or neighbouring plants (Spoel & Dong 2012). Beside the cell wall, that can be reinforced by callose, plants have two primary lines of defence (Spoel & Dong 2012).

1.2.1. PAMP triggered immunity

The first layer of defence consists of membrane standing immune receptors that recognize conserved structures of the different pathogens plants regularly encounter called pattern recognition receptors (PRR). This first layer of immunity is often labelled as pathogen-associated molecular pattern (PAMP) triggered immunity (PTI).

A regularly mentioned example of a PRR is the flaggelin sensitive 2 (FLS2) receptor from *A.thaliana* that recognizes bacterial flagellin (Boutrot & Zipfel 2017). Another example is the recognition of chitin, a substance found in the cell wall of fungi by the LysM containing receptor like kinase 5 (LYK5, Boutrot & Zipfel 2017). The receptors recognizing the pathogens are membrane bound receptor-like kinases or other receptor like proteins (Boutrot & Zipfel 2017). The extracellular domain that is responsible to bind the recognized pathogen associated structures, consists in most cases of a leucine-rich repeat (LRR) domain, alternatively lectin and LysM domains were identified (Boutrout & Zipfel 2017).

Most PRR identified in plants rely on the BRI1 associated receptor kinase 1 (BAK1), another receptor-like kinase with an extracellular LRR domain. Upon binding of the target molecule to the PRR, a dimer with BAK1 is formed that then leads to autophosphorylation of the receptors. Downstream signalling is usually achieved via other kinases (Dodds & Rathjen 2010). The subsequent immune response of the cell is described below, as the response is widely similar for both layers of the plant immune response (Dodds & Rathjen 2010).

1.2.2. Effector-triggered immunity

Successful pathogens found ways to overcome the PAMP-triggered immune response described above. This is achieved by inserting proteins in the plant cell that counter and suppress the immune response triggered by the activation of PRR. These molecules can consist of a wide variety of proteins and enzymes that can disrupt all layers of the plant immune response and are called effector proteins (Jones & Dangl 2006). Typically, each pathogenic bacterium injects 20-30 different effector proteins in the plant cell, eukaryotic pathogens can possess several hundred different effector proteins (Dodds & Rathjen 2010). Effectors can be inserted in the cell via the type III secretion system for bacteria, haustorium for fungi or stylets by insects (Dangl et al. 2013).

To defend against the effects of effector proteins, plants have developed a second layer of immunity called effector-triggered immunity (ETI). This layer of defence relies on a series of intracellular, soluble, receptors (Jones & Dangl 2006). These receptors consist of three domains: A nucleotide binding domain (NB), a leucin rich repeat (LRR) domain and either a coiled-coil (CC), coiled-coil RPW8-like (CC-R) or Toll-interleukin receptor like 1 domain (TIR, Jones et al. 2016).

Typically, the activation of these receptors leads to the controlled death of the cell so that the spreading of pathogens is limited (Jones & Dangl 2006). These receptors can either directly bind the effector proteins or guard a target protein (or a decoy) for the effects of effector protein, lastly some receptors are fused to additional domains that act as bait for effector proteins (Dodds & Rathjen 2010). The indirect detection of the presence of effector proteins allows the plant to use a single receptor to detect different pathogens. For example, the CC-NB-LRR protein HopZ activated resistance 1 (ZAR1) is able to confer resistance in response to seven effector proteins from three different pathogens (Guozhi et al. 2021). Additionally, some receptors work in pairs, as demonstrated by the receptor pair resistance to *Pseudomonas syringae* 4 (RPS4) and resistance to *Ralstonia solanaceranum* (RRS1) that form homo- and heterodimers. RPS4 recognizes the effector Avirulence RPS4 (AvrRPS4) from *P.syringae*, while RRS1 with its additional WRKY domain is responsible for detection of the effector *Pseudomonas* outer protein P2 (PopP2), an acetyltransferase from *Ralstonia solanacearum*. Additionally, both receptors are needed for defence against the pathogenic fungus *Colletotrichum higginsianum* (Narusaka et al. 2009).

Recent studies have greatly increased our understanding of how NLR receptor's function. A study involving cryo-electron microscopy (cryo-EM) revealed that upon activation, the CC-NB-LRR protein ZAR1 forms a pore in the membrane that allows for an influx of calcium ions (Guozhi et al. 2021). It is unknown if this a general mechanism or special for ZAR1. The TIR domain of TIR-NB-LRR receptors have been shown to catalyse the hydrolyzation of nicotinamide adenine dinucleotide (NAD^+) to adenosine diphosphate-ribose (ADP-ribose) and nicotinamide, additionally it has been proven that this activity is necessary for the appropriate immune response (Wan et al. 2019).

An additional group of NLRs is termed helper-NLR. They act downstream of the sensor-NLRs described above and are needed for a broad range of different receptors (Jubic et al. 2009). These helper NLRs have an atypical CC-domain with homology to the resistance to powdery mildew (RPW8) protein and are called RNL. In *A.thaliana*, two families with six isoforms are known and referred to as activated disease resistance 1 (ADR1) and N-requirement gene 1 (NRG1). Both protein families act downstream of the sensor-NLRs described above. NRG1 is essential for all tested TNL receptors to function, ADR1 is needed for correct immune response after activation of both CNL and TNL receptors (Jubic et al. 2019). ADR1 is highly conserved among all higher plants, and NRG1 is conserved among all dicots, showing the central function of these two proteins (Lapin et al. 2009). A schematic representation of the basic principles of the plant immune system is shown in figure 2.

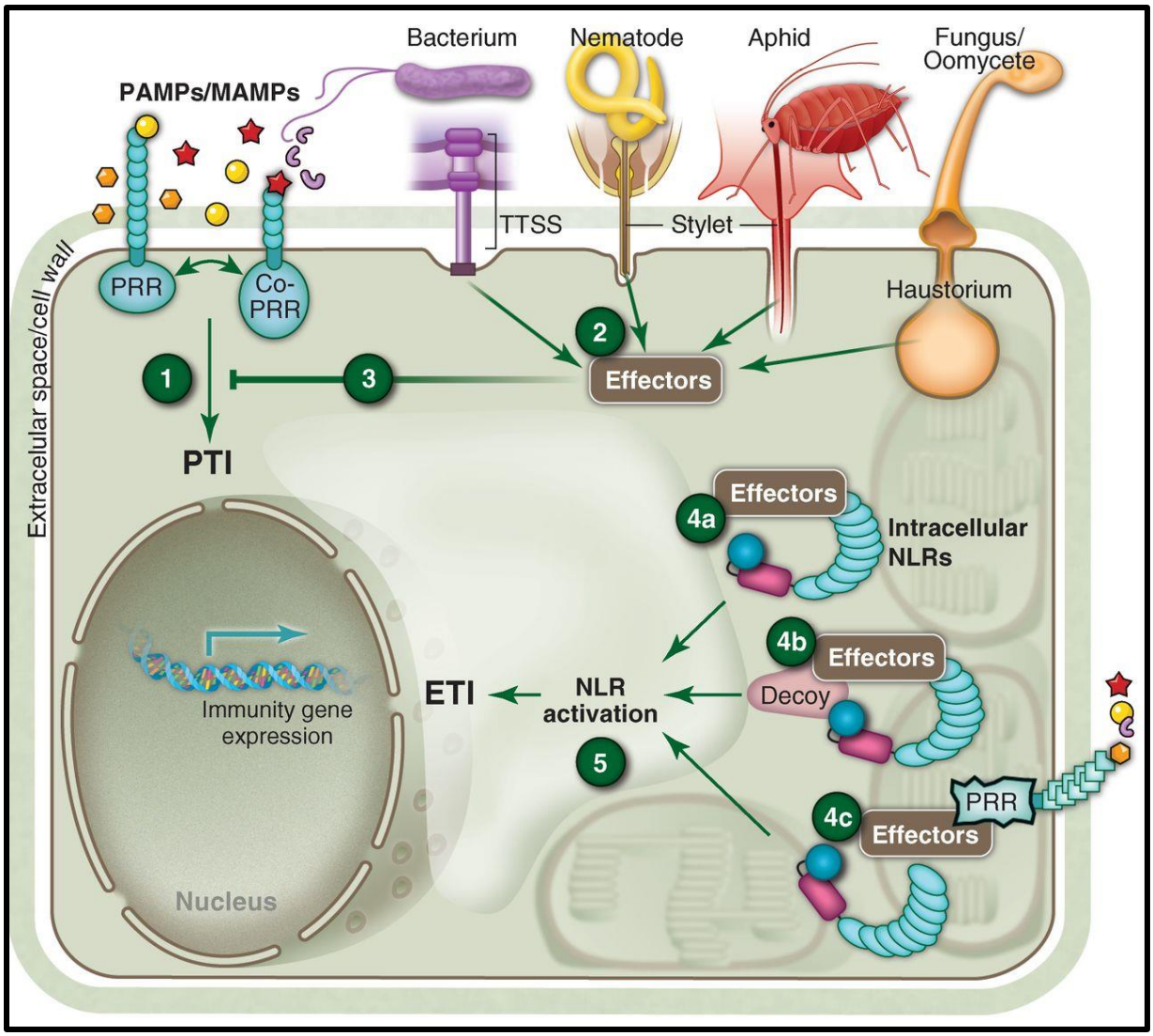

Figure 2: Schematic representation of the plant immune response. Membrane standing receptors recognize conserved molecular patterns and trigger the first layer of defence called pattern triggered immunity (1, PTI). Pathogens can overcome the PTI by injector effector proteins in the cell (2). These effector proteins can block the PTI (4). Effector proteins can be detected either directly or indirectly by the plant using cytoplasmatic receptors (4a-c). These receptors then trigger the second layer of defence that is called effector-triggered-immunity (5, ETI). Taken from Dangl et al. 2013 with friendly permission by AAAS.

1.2.3. The role of the EDS1 protein family in the plant immune defence

Beside the two helper-NLRs ADR1 and NRG1, which have been shown to be central points of the plant immune response (Jubic et al. 2009), the three members of the EDS1 protein family are known to be node points of the plant immune system (Dongus & Parker 2021). The EDS1 family consists of three members: EDS1, PAD4 and SAG101. EDS1 was first described in 1996 (Parker et al. 1996) followed by PAD4 1999 (Jirage et al. 1999) and SAG101 in 2002 (He & Gan 2002).

Both EDS1 and PAD4 are conserved among higher plants, while SAG101 is absent in monocots, indicating the importance of the EDS1 family proteins to allow the immune response in plants to function correctly (Baggs et al. 2020). A loss of EDS1 or PAD4 and to a lesser degree SAG101 leads to a dramatically increased susceptibility to a broad range of pathogens including bacteria, fungi and feeding insects (Wagner et al. 2013, Dongus et al. 2020, Feys et al. 2001, Feys et al. 2005).

All three members of the EDS1 protein family are confidently placed to act downstream of the immune receptors, but upstream of the actual immune response that will be discussed below. This means the three proteins are involved in the signalling chain from detection via receptor to immune response (Dongus & Parker 2021). The EDS1 signalling node is central for both ETI and PTI to function correctly (Dongus & Parker 2021). Nevertheless, the actual function of the three proteins is not yet known.

First, it should be mentioned that several interaction partners of EDS1 beside SAG101 and PAD4 were identified, usually using yeast-two-hybrid (Y2H) assays. As EDS1 is a central protein intricately involved in the immune response, it is a logical target for effector proteins inserted by pathogens to shut down the immune response. Three such proteins have been proposed to bind to atEDS1. HopA1 and AvrRPS4 are unrelated effector proteins from *P.syringae* and have both been shown to interact with EDS1 via bimolecular fluorescence complementation (BiFc) assays (Bhattacharjee et al. 2011); however, a direct physical interaction between EDS1 and AvrRPS4 is still discussed (Huh et al. 2017). Additionally, an effector from *P.capsici* has been recently shown to interact with EDS1 (Li et al. 2019).

Several cellular proteins have also been shown to bind EDS1. In the following overview, all further mentioned interaction partners refer to proteins from *A.thaliana*. EDR1 is a kinase that binds to both EDS1 and PAD4 and has been shown to interrupt the binding of the two proteins and acts as a negative regulator of EDS1 triggered immunity (Neubauer et al. 2020). The salicylic acid receptors NPR1-like protein 3 and 4 (NPR3 and NPR4) mediate the degradation of EDS1 by the proteasome and therefore also act as negative regulators of EDS1; *vice versa*, GH3-like defence gene 1 (PBS3) binds at the same site and protects EDS1 from being degraded (Chang et al. 2019). Another negative regulator of the EDS1 dependent immunity is repressor of RPS4-RLD 1 (SRFR1) that has been shown to be co-localized by BiFc assays (Bhattacharjeee et al. 2011). EDS1 directly binds to several DELLA proteins that are needed to regulate the immune response triggered by EDS1 (Li et al. 2019). Another negative regulator of EDS1 is EDS1-interacting J-protein 1 (EIJ1) a DnaJ-like protein that binds to atEDS1 and prevents its import in the nucleus (Li et al. 2021).

Lastly, EDS1 interacts with several different TNL receptors including RPS4, resistance to *Pseudomonas syringae* 6 (RPS6), variation in compound triggered root growth response (VICTR) and suppression of NPR1.1 (SNC1) as shown via BiFc (Heidrich et al. 2011, Bhattacharjee et al. 2011, Kim et al. 2012).

These findings show that EDS1, PAD4 and SAG101 are part of an intricate and complex network of different proteins that is responsible to activate and regulate the plant immune response.

Recent studies were able to clarify the role of the EDS1-SAG101 dimer. Analysis of the genome of different plants revealed that monocots and other plants lack both NRG1, a helper NLR, as well as TNLs (Lapin et al. 2020). Since the presence of atSAG101, atEDS1 and atNRG1 is able to recover the loss of these proteins in *N.benthamiana* and all three proteins act downstream of the immune receptors, it has become clear that these three proteins form a module required for TNL depended immunity (Lapin et al. 2019). The ability of TNL receptors to successfully trigger an immune response relies on the ability of the TIR-domain to hydrolyse NAD^+ to nicotinamide and ADP-ribose (Wan et al. 2019). While the reason for this catalytic activity is not yet clear, it is plausible to suggest a model where the products of the reaction mentioned above somehow activate the EDS1/SAG101/NRG1 module (Lapin et al. 2020). Another such module consists of the EDS1/PAD4 dimer as well the helper NLR ADR1. This module is similar to the EDS1/SAG101-NRG1 module essential for a successful TNL-dependent immunity, although it is not yet as deeply examined (Sun et al. 2021). The fact that the plant immune system relies on an intricate and complex number of interactions and crosstalk between the branches and pathways described above is demonstrated by the finding that the EDS1-PAD4-ADR1 module is needed for the PRR receptor protein suppressor of BIR1 (SOBIR1) to function correctly (Pruitt et al. 2021).

In conclusion it can be said that EDS1, SAG101 and PAD4 can be confidentially placed as central components of the cascade responsible to trigger the immune response after the immune receptors detect an invading pathogen. Despite the central role of these proteins their actual function is not clear as of now.

1.2.4. Response of the plant cells after detection of pathogens

After detection of the pathogen and subsequent signalling the individual plant cells can react with a multitude of defensive actions. The hormones salicylic acid and jasmonic acid accumulate in the plant cells after infection with a pathogen and are involved in the induction and regulation of the immune response (Wiermer et al. 2005). In general, salicylic acid mediated pathways are effective against biotrophs, whereas jasmonic acid mediated pathways are effective against necrotrophic pathogens and chewing insects (Wiermer et al. 2005). Plant cells can reinforce their cell walls, that act as a first physical barrier against invading pathogens by depositing callose after the activation of host defence pathways (Spoel & Dong 2012). Further effects consist of an influx of calcium ions, a burst of reactive oxygen species, and ultimately a reprogramming of the gene expression as well as the production of pathogenesis-related proteins (Dodds & Rathjen 2010).

These can include hydrolytic enzymes like chitinases or β-1,3-glucanases as well as anti-microbiological defensins and thaumatins (Spoel & Dong 2012). Ultimately, usually as effect of ETI, the localised controlled death of the cell is triggered, a process called hypersensitive response, that inhibits further growth of the invading pathogens (Dodds & Rathjen 2010). The whole plant can be primed against invading pathogens by signal molecules that are transported by the phloem through the plant and induce a broad-spectrum resistance among the whole organism that can include other pathogens and can last several days (Spoel & Dong 2012).

1.3. Structure biology of the EDS1 protein family

1.3.1. Structural knowledge of the EDS1 protein family

The main aim of this thesis is to broaden the knowledge of the structure of EDS1 and PAD4. The structure of the EDS1-SAG101 dimer from *A.thaliana* was solved via X-ray crystallography (Wagner et al. 2013) and was up to this work the only structural information available for the EDS1 protein family. A search in the PFAM database (Mistry et al. 2021), revealed the N-terminal domain of EDS1 to be classified as a lipase class 3 like domain. The central residues in the active centre of lipases with similar folds as the N-terminal domain of EDS1, PAD4 and SAG101 consists of a serine, a histidine and an aspartic acid, together called catalytic triad (Ollis et al. 1992). These three amino acids are conserved among species in EDS1 and PAD4, but not SAG101 (Wagner et al. 2013). Nevertheless, no catalytic activity was found for EDS1 *in vitro* and no phenotype was observed when mutating the catalytic triad of EDS1 (Wagner et al. 2013). While a heightened susceptibility to the green peach aphid (*Myzus persicae*) was found when mutating the catalytic serine or aspartic acid of PAD4, no such effect was observed when removing the central histidine residue of the catalytic triad, thus indicating that the observed phenotype is not the result of a loss of lipase like activity (Louis et al. 2012). Curiously, a lipase activity was reported in a study using recombinant SAG101 (He & Gan 2002) as the catalytic triad of SAG101 is, across the species, not present, this result cannot be the result of a canonical lipase-like activity (Wagner 2013). As mentioned above the N-terminal domain was found to be a α/β-hydrolase fold with the typical eight β-sheets surrounded by α-helices. Both EDS1 and PAD4 (in most species, see table 18) possess an additional insertion that is not part of the canonical fold (Wagner et al. 2013, Rauwerdink & Kazlauskas 2015). The atEDS1-atSAG101 structure revealed that in EDS1, this insertion blocks access to the active site and therefore explains the inactivity of EDS1 when tested for lipase-like activity (Wagner et al. 2013). Additionally, part of the potential oxyanion hole is blocked by a preformed hydrogen bridge and a phenylalanine residue is occupying the space of a potential substrate (Wagner et al. 2013). EDS1 forms mutually exclusive dimers with PAD4 and SAG101 (Wagner et al. 2013) and has been shown to form oligomers (Feys et al. 2001). The helix αH that is formed as part of the insertion is essential for the interaction of EDS1 with both PAD4 and SAG101 (Wagner et al. 2013).

After the EDS1/SAG101 crystal structure (Wagner et al., 2013) , it was attempted by several former members of our research group to produce crystals of unbound EDS1; such crystals could indeed be grown, but the resolution and general quality of their X-ray diffraction was never sufficient to solve a crystal structure (Wagner 2013; Christine Toelzer, unpublished results). To overcome this problem, two strategies were pursued: (i) A low-resolution small angle X-ray scattering (SAXS) structure of EDS1 was recorded by Christine Toelzer (Voss et al. 2019); (ii) the crystallization of unbound EDS1 should be – and this was the main task of this PhD project – improved by the help of nanobodies as described below in chapter 1.3.2

While EDS1 and SAG101 can be expressed easily in *E.coli*, up to now no report of the successful recombinant expression of PAD4 of any species is known. Despite numerous attempts to express atPAD4 in *E.coli* including the use of multiple strands of *E.coli*, expression of additional chaperones, refolding from inclusion bodies as well as several different solubility tags no soluble protein could be obtained until now (Wagner 2013, Klimpel 2014). Moreover, no report of the successful purification of any PAD4 orthologue is available until now.

1.3.2. Nanobodies as crystallisation chaperone for atEDS1

As described above, crystals formed by unbound EDS1 show poor diffraction quality insufficient for structure determination. As classical attempts to improve the diffraction quality had failed (Wagner 2013; Christine Toelzer, unpublished results), in a next step nanobodies were created to trap unbound EDS1 in its potentially specific conformational state and to obtain crystals enabling to determine the structure to at least medium resolution.

Nanobodies are antibody fragments from either *Camelidae* or sharks that can be easily expressed in *E.coli* (Muyldermans 2013). While common mammalian immunoglobulin G (IgG) antibodies are made from four chains that are linked together via cysteine-bridges, with two chains making up the specificity determining region (Figure 3A), *Camelidae* and sharks possess in addition special antibody that are made from just two proteins (Figure 3B+D). The antigen binding region of these antibodies is made up by a single protein domain and is called nanobody when expressed separately (Muyldermans 2013). A representation of the different mentioned antibody types is shown in figure 3.

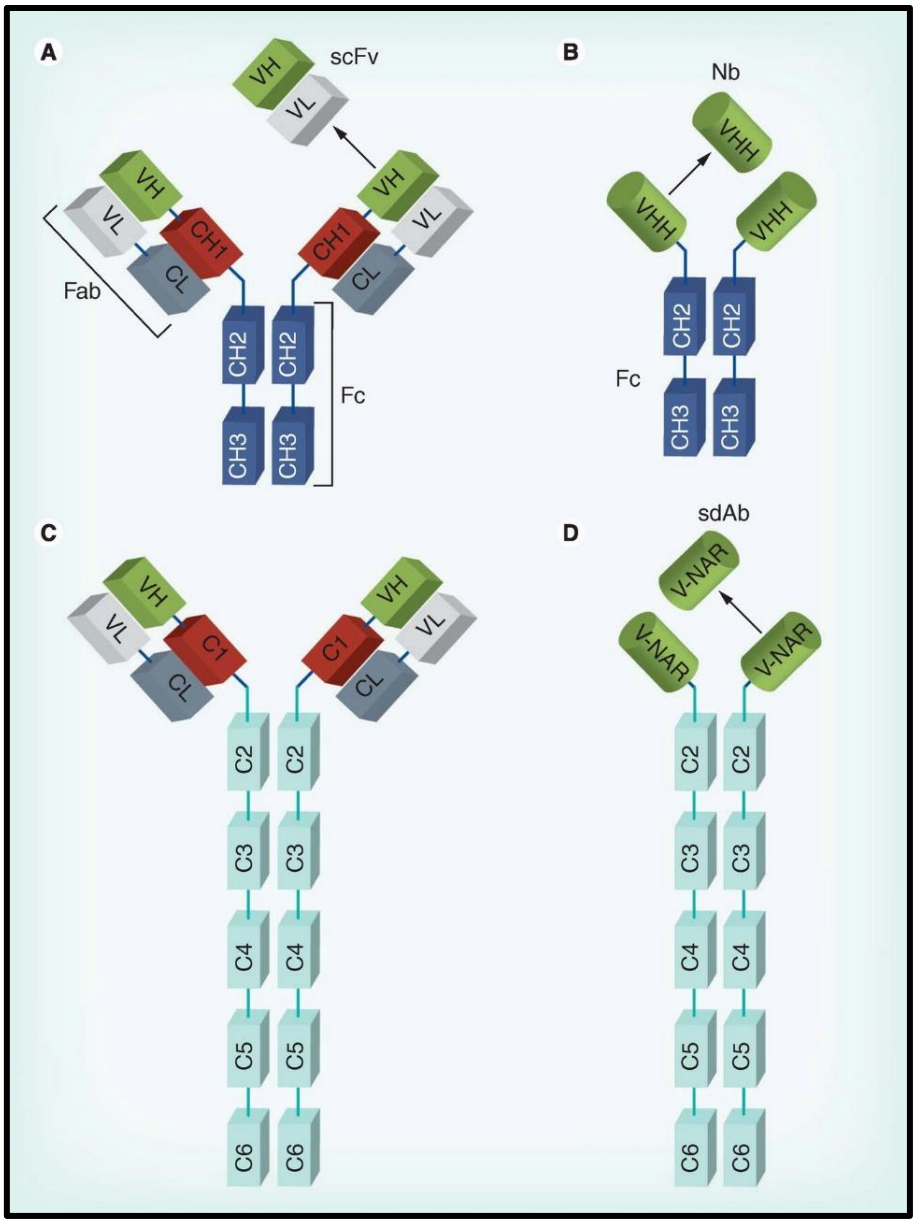

Figure 3: Schematic representation of different types of antibodies from Camelidae, sharks and cartilaginous fish. Figure A: common IgG found in mammal, consisting of four chains linked via cysteine bridges. The smallest antigen binding entity consists of two chains and is shown separately. Figure B: Heavy chain only IgG found in Camelidae that consist of two protein chains linked via cysteine bridges. The smallest epitope binding entity is shown separately and consists of only one protein. Figure C: IgW found in cartilaginous fish. This antibody class, similar to mammalian IgG consists of four chains, linked by disulphide bridges. The smallest epitope binding entity consists of two protein chains. Figure D: Ig-NAR as found in sharks, similar to camelid heavy chain only antibodies that consist of two protein chain. Again, the smallest epitope binding entity is made up by a single protein chain. VH= Variable part of heavy chain. VL= Variable part of light chain. CH= Constant part of heavy chain. CL= Constant part of light chain. scFv= Single chain variable fragment. Nb= Nanobody. V-NAR= variable new antigen receptor. sdAb= single domain antibody fragment. Taken from Hassanzadeh-Ghassabeh et al. 2013 with friendly permission of Future Medicine LTD.

Nanobodies are small, highly stable in a broad range of conditions, can be easily expressed in high yields and all kind of expression systems, are highly soluble, can be easily refolded, typically have subnanomolar affinities to their antigens and are not immunogenic. All these characteristics make nanobodies highly attractive as a tool for various biotechnological, therapeutical and diagnostical purposes (Hassanzadeh-Ghassabeh et al. 2013).

Nanobodies are classically generated by injecting the animal with highly purified protein several times over a period of several weeks. In this time the animal immune system produces specific antibodies with high affinities due to the somatic hypermutation that occurs during the maturation of the B-cells. A small amount of blood is then taken from the animal and anticoagulants are added. Afterwards the lymphocytes are isolated, and the RNA content of the cell is extracted. The messenger-RNA (mRNA) is then transcribed to complementary-DNA (cDNA) using oligo dT primer. The antigen binding entity of the heavy chain only antibodies are amplified via polymerase chain reaction (PCR) and then cloned in a vector suitable for phage-display followed by several round of selection against the target protein (Muyldermans 2013). The resulting sequences of interest can then easily be expressed in a broad range of expression systems (Muyldermans 2013). An alternative approach consists of the use of a library of randomly generated nanobodies displayed by yeast cells (McMahon et al. 2018). This method negates the use of an animal as well as the necessity to produce protein to inject into the animal, a distinct advantage for proteins hard to express in large amounts.

The first drug based on nanobodies, a protein binding the von-Willebrandt factor, was approved in 2018 for treatment of thrombotic thrombocytopenic purpura[5], a rare disease that leads to the spontaneous formation of thrombi. Another proposed application is the use of a trimeric nanobody against the SARS-CoV2 spike protein in an aerosol to prevent infection (Schoof et al. 2020). A similar study, this time using shark-derived sequences was recently published (Gauhar et al. 2021). A different, research-based application of nanobodies is the use as crystallization chaperone. Nanobodies can trap proteins in a certain conformation, stabilize the crystal packing and subsequently the diffraction quality of crystals (Traenkle et al. 2016). The use of nanobodies as chaperones for crystallisation allowed the structure determination of several challenging targets including G-protein coupled receptors, transporters and other membrane proteins (Traenkle et al. 2016).

5 https://www.ablynx.com/rd-portfolio/clinical-programmes/caplacizumab/

The impact of these studies is highlighted by the awarding of the Nobel prize for chemistry 2012 to Brian Kobilka and Robert Lefkowitz for their work on G-protein coupled receptors. Kobilka and colleagues published several studies of high impact concerning nanobodies and the β_2-adrenoreceptor (Westfield et al. 2011, Rasmussen et al. 2011 (1), Rosenbaum et al. 2011, Rasmussen et al. 2011 (2)), additionally the use of nanobodies is highlighted in the published scientific background put together by the royal Swedish academy concerning the awarding of the Nobel prize for chemistry 2012[6].

1.4. Thesis aims

As shown above, the members of the EDS1 protein family are central proteins of the plant immune system. Nevertheless, central question about both the structure and the function of EDS1, PAD4 and SAG101 remain unanswered. This thesis aim is to broaden the knowledge about the structure of the EDS1-protein family.

The first question that I will attempt to answer is how the structure of unbound atEDS1 deviates from the atEDS1-atSAG101 complex. In this context nanobodies will be used as both crystallisation chaperones and as a tool to trap a potential alternative conformation. Additionally, I aim to answer the question if atEDS1 is in its unbound form really a dimer as is believed today (Feys et al. 2005). The method of choice to answer the questions raised above will be x-ray crystallography.

The next aim of this thesis is to recombinantly express and purify a PAD4 variant in *E.coli.* For this task two avenues seem plausible: Using a different expression system or using an orthologue of PAD4. In this thesis I will use sequences of PAD4 orthologues from several different species to increase the chances to gain soluble PAD4 in amounts suitable for crystallisation experiments. The ultimate aim will be to assemble an EDS1-PAD4 complex for characterisation and structure determination.

[6] https://www.nobelprize.org/uploads/2018/06/advanced-chemistryprize2012.pdf

2. Materials and methods

2.1. Materials

2.1.1. Chemicals

All chemicals used were purchased in laboratory grade quality from Sigma Aldrich, Roth, Alfa Aesar, Thermo scientific or Applichem if not otherwise stated.

2.1.2. Enzymes and proteins

Table 1: Proteins and enzymes used in this thesis.

Enzyme	**Supplier**
DNaseI	Roche
Lysozyme from chicken egg white	Sigma Aldrich
BsaI	New England biolabs
DpnI	New England biolabs
FseI	New England biolabs
AatII	New England biolabs
SalI-HF	New England biolabs
NotI-HF	New England biolabs
T4 DNA polymerase	Thermo fisher
Mucor miehei lipase	Sigma Aldrich
Ferritin from horse spleen	Sigma Aldrich
Lactate dehydrogenase from porcine heart	Sigma Aldrich
Bovine serum albumin	Sigma Aldrich
Pfu-Polymerase	Expressed and purified in Baumann lab (institute of biochemistry, University of cologne)
Fast alkaline phosphatase	Thermo fisher
T4 DNA Ligase	Thermo fisher

All enzymes and proteins were stored at either 4°C or -20°C as specified by the supplier and were thawed freshly before use.

2.1.3. Oligonucleotides

All oligonucleotides were purchased from Sigma Aldrich in desalted quality. The lyophilised primer were solved in water to a final concentration of 100 µM.

2.1.4. Buffers and solutions

The used buffers and solutions are listed below in table 2. All buffers and solution were prepared with water desalted and filtered by a Puranity™, ultrapure water filtration system (VWR). All buffers and solutions were filtered and degassed using a 0.45 µM filter (Mixed cellulose ester, Fisher Scientific) and a vacuum pump. All pH values were adjusted by addition of either hydrochloric acid or sodium hydroxide. Phosphate buffers were created by mixing equally concentrated solutions of K_2HPO_4 and KH_2PO_4 until the desired pH was reached. Dithiothreitol (DTT) was added fresh just before use from a 1M stock solution stored at -20 °C.

Table 2: Buffers and solutions used in this thesis.

DNA and protein electrophoresis	
50x Tris-acetate-EDTA (TAE) buffer	50 mM EDTA, 2M Tris, 1M acetic acid
Coomassie staining solution	40% (v/v) ethanol, 10% (v/v) acetic acid, 0.1% Coomassie brilliant blue
10x Running buffer	250 mM Tris, 1.92 M glycine, 1% (w/v) SDS
6x Sample buffer	60 mM Tris-HCl, 12% (w/v) SDS, 47% (v/v) glycerol, 20 mM DTT, 0.6% (w/v) Bromophenol blue, pH=6.8
Stack buffer	1 M Tris pH=6.8
Separationgel buffer	1.5M Tris/HCl, pH= 8.8
Destaining solution	20% (v/v) ethanol, 5% (v/v) acetic acid
PCR	
10x PCR reaction buffer	500 mM KCl, 100mM Tris, 15 mM $MgCl_2$, 1% (v/v) Triton X-100, pH=9
dNTP mix	2.5 mM dATP, 2.5 mM dCTP, 2.5 mM dGTP, 2.5 mM dTTP
Protein purification	
atEDS1 Lysis/Wash buffer	50 mM HEPES, 50 mM NaCl, 10% glycerol, 20 mM imidazole, 1mM DTT pH=8
atEDS1 Elution buffer	50 mM HEPES, 50 mM NaCl, 250 mM imidazole, 1mM DTT, pH=8
atEDS1 Gelfiltration buffer	50 mM HEPES, 50 mM NaCl, 1% glycerol, 1 mM DTT, pH=8
vvEDS1/vvPAD4 Lysis/Wash buffer	50 mM HEPES, 300 mM NaCl, 10% glycerol, 20 mM imidazole, 1mM DTT pH=8

vvEDS1/vvPAD4 Elution buffer	50 mM HEPES, 300 mM NaCl, 250 mM imidazole, 1mM DTT, pH=8
vvEDS1/vvPAD4 Gelfiltration buffer	50 mM HEPES, 300 mM NaI, 1% Glycerol, 1 mM DTT, pH=8
Nanobody wash buffer	100 mM potassiumphosphate, 20 mM imidazole, pH=7
Nanobody elution buffer	100 mM potassiumphosphate, 250 mM imidazole, pH=7
Immunoblotting	
Tris buffered saline with Tween-20 (TBST)	10 mM Tris, 150 mM NaCl, 0,05% Tween 20, pH=7.5
Towbin buffer	25 mM Tris, 0.192 M glycine, 1% (w/v) SDS, 10% methanol (v/v)
Enhanced chemiluminescence lightning solution	666 µl separation gel buffer, 25 µl 90 mM p-coumaric acid, 50 µl 250 mM luminol, 3 µl 30% H_2O_2
Other buffers and solutions	
Inoue buffer	10 mM PIPES, 250 mM KCl, 15 mM $CaCl_2$, 55 mM $MnCl_2$, pH=6.7 adjusted with KOH

2.1.5. Media for bacterial growth

All media were either freshly prepared or autoclaved before use. All media were prepared using tap-water. To prepare agar-plates 15 g/l agar was added before autoclaving. For Terrific broth (TB) medium the phosphates were prepared as 10x stock solution and autoclaved separately. For superoptimal broth medium with catabolite repression (SOC) medium the glucose was added after autoclaving.

Table 3: Media for bacterial growth used in this thesis.

Lysogeny Broth (LB)-medium	10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl
TB-medium	10 g/l tryptone, 20 g/l yeast extract, 5 g/l glycerol, 2.3 g/l KH_2PO_4, 12.5 g/l K_2HPO_4
SOC-medium	20 g/l tryptone, 5 g/l yeast extract, 0.5 g/l NaCl, 0.2 g/l KCl, 0.9 g/l $MgCl_2$, 1.2 g/l $MgSO_4$, 3.7 g/l Glucose
Low salt LB-medium	10g/l tryptone, 5 g/l yeast extract, 0.5 g/l NaCl

2.1.6. Chromatographical matrices

Table 4: Chromatographical matrices used in this thesis.

Superdex200 10/300 GL	GE Healthcare
HiLoad Superdex75 16/60 prepgrade	GE Healthcare
NiNTA agarose	Cube Biotech
HisTrap ff 1ml	Ge Healthcare
Superdex200 16/60 prepgrade	GE Healthcare

2.1.7. Antibiotics

Table 5: Antibiotics used in this thesis.

1000x Kanamycin	50 mg/ml
1000x Ampicillin	100 mg/ml in 50% ethanol
1000x Chloramphenicol	50 mg/ml in ethanol
1000x Gentamicin	20 mg/ml
1000x Streptomycin	100 mg/ml
1000x Tetracycline	10 mg/ml in ethanol

2.1.8. *E.coli* strains

Table 6: E.coli strains and their genotypes used in this thesis.

Name	**Supplier**	**Genotype**
BL21(DE3)	Invitrogen	F^- *ompT* $hsdS_B$ (r_B–, m_B–) *gal dcm* (DE3)
DH5α	Invitrogen	F^- *ompT* $hsdS_B$ (r_B–, m_B–) *gal dcm* (DE3)
Rosetta™ (DE3)	Novagen	F^- *ompT* $hsdS_B$(r_B^- m_B^-) *gal dcm* (DE3) pRARE (Cam^R)
Rosetta™ (DE3) pLysS	Novagen	F^- *ompT* $hsdS_B$(r_B^- m_B^-) *gal dcm* (DE3) pLysSRARE (Cam^R)
BL21(DE3)pLysS		F^-, *ompT*, *hsdS*$_B$ (r_B–, m_B–), *dcm*, *gal*, λ(DE3), pLysS, Cm^r.
XL1-blue	Agilent	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F′ proABlacIqZΔM15 Tn10 (Tetr)].
BL21(DE3)-GenX	Genlantis	
WK6	VIB	F' lacIq delta(lacZ)M15 proA+B+ delta(lacproAB) galE rpsL
Arctic Express (DE3)	Agilent	E. coli B F– ompT hsdS(rB– mB–) dcm+Tetrgalλ(DE3) endA Hte [cpn10 cpn60 Gentr]

2.1.9. Plasmids and constructs

2.1.9.1. Plasmids for additional chaperones

To improve the solubility of proteins *E.coli* BL21(DE3) cells were transformed with an additional plasmid before preparation of competent cells. Plasmids were purchased from Takara Bio Inc.

Table 7: Plasmids encoding for additional chaperones to increase solubility of recombinantly expressed proteins. The plasmids, that express a chloramphenicol resistance were transformed in E.coli BL21(DE3) cells in addition to the expression plasmids.

Name	**Plasmid**	**Expressed chaperone**
Takara 1	pG-KJE8	dnaK-dnaJ-grpE- groES-groEL
Takara2	pGro7	groES-groEL
Takara3	pKJE7	dnaK-dnaJ-grpE
Takara4	pG-Tf2	groES-groEL-tig
Takara5	pTf16	tig

2.1.9.2. Plasmids and constructs used in this work.

Table 8: Plasmids for protein expression used in this thesis. All plasmids encode for a His-tagged protein.

pET22a-atEDS1	Stephan Wagner (AG Karsten Niefind)
pET42a-stPAD4	Stephan Wagner (AG Karsten Niefind)
pETM11-atPAD4	Christine Tölzer (AG Karsten Niefind)
pET42a-aaPAD4	Stephan Wagner (AG Karsten Niefind)
pETM11-atPAD4 S118A	This work/Alexander Rothemann
pETM11-vvPAD4	This work
pETM11-osPAD4	This work
pETM11-hvPAD4	This work
pETM11-vvEDS1^{Nterm}	This work
pETM11-vaEDL2	This work
pETM11-vvPAD4S123A	This work
pETM11-vvPAD4H305A	This work
pRSFDuet-atEDS1+atPAD4	Stephan Wagner
pET42a-EDS1(1-384)	Stephan Wagner

2.1.9.3. EDS1 specific nanobodies

All nanobodies were delivered in the pMECS vector by NSF-VIB.

Table 9: atEDS1 specific nanobodies used in this thesis. Members of the same group bind the same epitope but show slight differences in the backbone of the protein.

Group	**Member(s)**
1	1AT80, 2AT3
2	1AT3
3	1AT23
4	1AT55, 1AT73, 2AT17, 2AT8
5	1AT15, 1AT22, 1AT85, 1AT29
6	2AT2
7	1AT59
8	2AT11, 2AT19, 2AT13, 1AT66, 1AT45, 1AT94, 2AT22
9	1AT5, 1AT67
10	1AT1, 2AT16, 2AT18, 1AT38
11	1AT74
12	1AT13
13	1AT27
14	1AT20, 1AT72, 2AT6, 1AT78, 2AT9
15	1AT61, 1AT84
16	2AT24
17	1AT21, 2AT7, 1AT81
18	1AT86
19	2AT4
20	2AT20
21	1AT32
22	2AT26

2.1.10. Crystallization screens

Table 10: Commercially available screens for crystallization. Each screen contained 96 different conditions.

Name	**Manufacturer**
PEG/Ion HT	Hampton Research
PACT++	Jena Bioscience
Index HT	Hampton Research
Wizard I&II	Jena Bioscience
Crystal HT	Hampton Research
JCSG+	Molecular Dimensions
MIDAS HT	Molecular Dimensions
Morpheus	Molecular Dimensions
SaltRX	Home made by AG Prof. Dr. Ulrich Baumann
Ligand friendly screen	Molecular Dimensions
ProPlex	Molecular Dimensions

2.2. Methods

2.2.1. Molecular biological methods

2.2.1.1. Amplification of DNA via PCR

DNA fragments were amplified using PCR (Mullis et al. 1986). Primer flanking the desired sequence complementary to the sense and anti-sense strand were designed and synthesized chemically. Repetitions of DNA denaturing by heat, followed by cooling to anneal the primers and heating to reach the optimum of the thermostable polymerase, allowed exponential amplification of the desired DNA sequence. The typical temperature profile as well as reaction mixture are shown in table 11.

Table 11: Reaction mixture used for PCR reactions.

Reaction mixture	
10x PCR reaction buffer	5 µl
dNTP mix	2 µl
Forward primer (100 µM)	1 µl
Reverse primer (100 µM)	1 µl
Pfu DNA Polymerase	2 µl
Template DNA	0.5 µl
Water	38,5 µl

Table 12: Temperature profile used for PCR reactions.

Stage	Temperature (°C)	Time	No. of cycles
Initial denaturation	95	5 min	1
Denaturation	95	1 min	30-35
Annealing	55-68	1 min	
Extending	68	1 min per 500 bp	
Final extension	68	10 min	1

2.2.1.2. Site directed mutagenesis.

To introduce specific mutations a site directed mutagenesis PCR was performed. For that, complementary primers containing the desired mutation were designed, using the QuickChange primer design online tool by Agilent[7]. Afterwards a PCR was performed amplifying the whole plasmid, leading to a linear accumulation of the desired product containing the mutation. For the PCR the same protocol as shown above in table 12 was used with 55 °C as annealing temperature and eighteen cycles of amplification. As both the template plasmid as well as the linear product are present after the PCR, a DpnI digest was performed to remove the original plasmid. DpnI is a restriction enzyme that digests methylated DNA, methylation of the DNA occurs in *E.coli* cells.

2.2.1.3. Agarose gel electrophoresis of DNA

Due to the negatively charged phosphates in the backbone of nucleic acids, DNA migrates to the anode in an electric field. In an agarose gel matrix, they can therefore be separated by size. To create such a gel 1% (w/v) agarose in TAE buffer was boiled until the solution was clear. After the solution cooled down to roughly 50 °C, Midori Green (Nippon genetics) was added to label the DNA as per the manufacturer's specifications. Afterwards, the agarose was poured in form and left to cool and harden. The electrophoresis was performed in TAE buffer. Samples were mixed with 6x DNA loading buffer before application. To determine the size of the fragments a standard (1kbp marker New England biolabs) was applied in a separate lane. Electrophoresis was run at 200V, the gel was evaluated afterwards under UV light.

7 https://www.agilent.com/store/primerDesignProgram.jsp

2.2.1.4. Purification of DNA fragments from agarose gels

DNA fragments were illuminated under Ultraviolet (UV) light and excised from an agarose gel using a clean scalpel. To remove the agarose and clean up the DNA a commercially available kit (GenElute™ Gel Extraction Kit, Sigma-Aldrich) was used as per the manufacturer's instructions. The eluted DNA was stored at -20 °C.

2.2.1.5. Quantification of DNA solutions

The concentration of DNA solutions was determined using photometry, as DNA possesses an absorption maximum at 260 nm. 1 µl of the solution in question was applied to a photometer to measure an absorption spectrum after blanking the instrument using elution buffer or water. The concentration was then determined using Lambert-Beer's law and the extinction coefficient of double stranded DNA of 0.02 $(\mu g/ml)^{-1}\ cm^{-1}$. Additionally, the ratio of the A_{260}/A_{280} was determined to evaluate the purity of the DNA solution. A pure solution of double stranded DNA is expected to have a ratio of 1.8, lower values point towards protein impurities, and a higher value points towards other contaminants like phenols.

2.2.1.6. Preparation of plasmids from E.coli cultures

5 ml of LB medium with the appropriate antibiotics was inoculated with a single colony from a selective agar plate. The culture was then incubated at 37 °C and 180 rpm in a shaker overnight. The next day the cells were harvested by centrifugation at 13000 xg for one minute. The plasmids were isolated using a commercially available kit (GenElute™ Plasmid Extraction Kit, Sigma-Aldrich) as per the manufacturer's instruction. The kit is based on the alkaline lysis method. The eluted DNA was stored at -20 °C.

2.2.1.7. Cloning via restriction digestion

The desired insert was amplified via PCR using primers that contained the recognition sequence of the desired restriction enzymes. Restriction enzymes were selected to not cut the insert, and the vector just once in the multiple cloning site. The PCR product was mixed with the appropriate buffer as supplied by the manufacturer and 1 µl of both restriction enzymes. The mixture was incubated, depending on the specific enzyme, for 1h or overnight, at 37 °C. The vector was treated similarly, but 1 µl of fast alkaline phosphatase was added additionally. Both the PCR product and the vector were then purified via agarose gel electrophoresis as described above. Afterwards 25 ng vector and product in a molar relation of 1:3, 1:5 and 1:10 were mixed with 2 µl ligase buffer (provided by the manufacturer), 0.5 µl T4 DNA Ligase and water to a volume of 20 µl. Additionally a sample without insert was prepared to control the successful digestion. The mixtures were incubated overnight at room temperature. Afterwards competent *E.coli* were transformed with the ligated plasmids and a colony PCR was performed to verify the presence of the target.

2.2.1.8. Ligase independent cloning

The vectors and instructions were provided gracefully by the EMBL Hamburg. All cloning was done as per the instructions provided [8]. In short, 5µg of the vector was purified and digested with BsaI and purified again. Afterwards the vector was treated with T4 DNA polymerase in the presence of dTTP to create a single stranded overhang. Hereby the 3´->5´ exocuclease function of the polymerase is used to remove bases until the first adenine. The insert was amplified using primers with the specified overhang. After the PCR the insert was purified, 0.2 pmol of the insert were then treated with T4 DNA polymerase and dATP. Both reactions were stopped, using heating to 75 °C for 20 min. Afterwards vector and insert were mixed and after 10-minute incubation transformed in competent cells. The success was evaluated using colony PCR to verify the presence of the insert.

2.2.1.9. Colony PCR

To verify the presence of the insert after cloning, a colony PCR was performed. Hereby whole cells from an agar plate were used as template instead of isolated DNA. Concrete a small amount of a single colony from an agar plate was added to the PCR mixture as described in table 11 using half the volumes. Afterwards the result was analysed in an agarose gel and positive clones, showing the expected fragment, were further analysed by sequencing the purified plasmids.

2.2.1.10. DNA sequencing

To guarantee the correct insertion of the gene and to exclude that mutations occurred during the PCR, all plasmids were sequenced. For that, purified plasmids and fitting primers were mixed and send to the GATC sequencing service. Alternatively, a PCR was performed and the purified product sequenced. The results were analysed using SnapGene viewer (GSL Biotech LLC).

2.2.1.11. Preparation of chemically competent E.coli

5 ml of LB medium was inoculated with the desired *E.coli* strain and incubated overnight at 37 °C and 180 rpm. The next day 100 ml low salt LB medium was inoculated with the overnight culture and grown at 30 °C until an OD_{600} of 0.45 was reached. The flask was then cooled in an ice-water bath for twenty minutes. The cells were precipitated in a precooled centrifuge at 2500xg for 10 minutes. The cells were then resuspended in 32 ml ice-cold Inoue solution and again precipitated at 2500xg for 10 minutes. The pellet was resuspended in 8ml ice-cold Inoue solution and 600 µl warm dimethyl sulfoxide was added afterwards. The cells were incubated for 10 minutes on ice and then frozen in liquid nitrogen in precooled tubes in 100 µl aliquots. The cells were then stored at -80°C until use.

8 https://www.helmholtz-muenchen.de/fileadmin/PEPF/Protocols/LIC-cloning.pdf

2.2.1.12. Transformation of chemically competent E.coli cells

An aliquot of chemically competent cells was thawed on ice before either 0.5-1 µl of purified plasmid or 5 µl ligation mixture, were added. After a 20-minute incubation on ice the cells were heat shocked for 90 seconds at 42 °C. Directly afterwards 300 µl of SOC medium was added and the cells incubated at 37 °C for sixty minutes. The cells were then spread on an LB-agar plate with the appropriate antibiotics. The plates were incubated at 37 °C overnight and stored sealed at 4°C.

2.2.2. Protein biochemical methods

2.2.2.1. Test expression of recombinant proteins

To test the expression level and solubility of recombinant proteins the appropriate plasmid was firstly transformed in expression strains of *E.coli.* Afterwards 5 ml LB-medium with the fitting antibiotics was inoculated with a single colony using a pipette tip. The culture was then incubated at 37 °C and 180 rpm overnight. The next day, 200 ml autoclaved LB-medium was inoculated with the overnight—culture and incubated at 37° C until an OD_{600} of 0.6 was reached. If plasmids expressing additional chaperones were included, arabinose and/or tetracycline was added to the medium to induce expression of the chaperones. Afterwards 100 µl of the culture was removed and centrifuged at 11000xg for 1 min to precipitate the cells. The supernatant was removed, and the cells stored at -20 °C. The protein expression was then induced by the addition of 0.5 mM isopropyl-β-D-thiogalactopyranosid (IPTG), a non-hydrolysable activator of the lac-operon. All plasmids used contain a lac-operator downstream of the sequence of the target gene, as well as an additional gene for the lac-inhibitor, therefore the expression of the target protein can be induced by adding lactose or IPTG. The culture was then incubated at 18 °C and 180 rpm overnight. The next day again 100 µl of the culture were taken aside, precipitated and stored at -20°C, while the rest of the cells were harvested by centrifugation at 5000xg and 4 °C for 20 minutes. The cells were stored afterwards at -20 °C. Later the cells were thawed on ice for one hour. They were then resuspended in 15 ml of cold lysis/wash buffer supplied with 0.8 mg/ml lysozyme, 10 µg/ml DNaseI and 0.05% (v/v) triton-X100. Lysozyme is an enzyme and part of the immune system, it catalyses the hydrolysation of the 1,4-beta linkage of the N-acetylmuramic acid and the N-acetyl-D-glucosamine backbone of the bacterial cell wall. After thirty minutes the cells were lysed using sonification for 1 minute at 30% amplitude. The debris was then removed via centrifugation at 20.000xg for thirty minutes at 4 °C. A sample was taken from both the pellet and the supernatant for further analysis. All samples were treated with sample buffer and analysed using SDS gel electrophoresis and Coomassie staining.

2.2.2.2. Large scale expression of recombinant proteins.

An overnight culture of 100 ml LB-medium was inoculated with a single colony from a fresh agar-plate and incubated over night at 37 °C and 180 rpm. The next day 500 ml LB or TB-medium with the appropriate antibiotics was filled in a 2l baffled flask and inoculated with 5 ml of the overnight culture. The cells were grown at 37 °C until an OD_{600} of 0.6 (LB-medium) or 2 (TB medium) was reached. The expression of the recombinant protein was induced by adding 0.5 mM IPTG to the cells, that were then incubated overnight at 18 °C or 28 ° in the case of the nanobodies. The next day the cells were harvested by centrifugation at 3300xg and 4 °C for 30 minutes. The cells were resuspended in 0.9 % (w/v) NaCl and once again precipitated by centrifugation at 2500xg for 20 minutes. The cells were either used directly or stored at -80°C.

2.2.2.3. Purification of nanobodies.

The pMECS vector, used for overexpression of the nanobodies, encoded for a C-terminal HA- and His-tag, as well as a N-terminal pelB leader sequence. This sequence leads to the relocation of the freshly translated proteins in the periplasm, where the cysteine bridges are correctly formed, and the signal peptide is cleaved. Therefore, a complete lysis of the cells is not necessary, as just the outer membrane needs to be removed. In detail, cells from the -80 °C freezer were thawed at room temperature for thirty minutes. The thawed cells were then resuspended in 12 ml TES buffer per litre cell culture and shaken for one hour at room temperature. Afterwards, 18 ml per litre cell culture of 1:4 with water diluted TES buffer were added and the cells incubated for another hour. The osmotic shock, due to the sudden decrease in ionic strength leads to breaks in the outer membrane and the release of the periplasm in the supernatant. The remaining protoplast were then removed via centrifugation at 15000xg for 45 minutes. One millilitre Ni-NTA agarose was then added to the cleared supernatant to bind the His-tagged nanobodies. The suspension was incubated for one hour at room temperature to guarantee complete binding of the tagged protein to the matrix. Afterwards the matrix was removed from the supernatant in a gravity flow column, the beads were then washed three times with 15 ml nanobody wash buffer. After the last washing step 5 ml elution buffer were added and caught in a tube after 5 minutes incubation. An additional elution step was performed overnight. The eluted protein was concentrated using a centrifugal filter unit with a 15 kDa cut-off via. Afterwards, using the same filtration device the buffer was exchanged to gelfiltration buffer in several steps. The protein was stored at 4 °C. For long-time storage 0,05% NaN_3 was added to prevent growth of microorganisms.

2.2.2.4. Purification of His-tagged proteins.

Cells were taken from the freezer and thawed on ice for one hour. Afterwards the cells were resuspended in cold Lysis/Wash buffer with addition of 0.8 mg/ml lysozyme, 10 µg/ml DNaseI and 0.05 % triton-X100 until a volume of 50 ml per 2l medium was reached. The cells were incubated on ice for one hour to guarantee a complete lysozyme digestion. The cells were then lysed using sonification on with 30% amplitude for 2 minutes with two seconds break after a two second pulse to guarantee an optimal cooling of the cells. Afterwards the cell debris was precipitated using centrifugation for 20000xg for 45 minutes at 4 °C. The supernatant was applied to a 1ml NiNTA cartridge using a 50 ml superloop (GE Healthcare) via an Äkta Purifier/Äkta Explorer/ Äkta PrimePlus (GE Healthcare/Amersham). The flowrate was set to 0.5 ml/minute to guarantee an effective binding of the proteins to the matrix. Afterwards the column was washed with 20 ml Lysis/Wash buffer and the protein was eluted with a gradient from 0-100% elution buffer. The eluate was collected in fractions of 1.5 ml. During the whole chromatography, the absorption at 280 nm was measured, as the aromatic amino acids, especially tryptophane, of proteins do have an absorption maximum at that wavelength. Fractions containing protein, as observed by the absorption at 280 nm, were pooled and concentrated using a 30 kDa cut-off filtration device. The protein was always stored on ice or in the fridge. Proteins were used for further analysis immediately or the next day.

2.2.2.5. Size exclusion chromatography

To further purify proteins after a Ni-NTA affinity chromatography, a size exclusion chromatography was performed. A size exclusion chromatography separates particles depending on their hydrodynamic radius. A column is filled with porous material that allows smaller molecules to enter the pores, therefore these smaller particles have a longer way until they reach the bottom of the column. For preparative purifications a Superdex75 16/60 with a column volume of 120 ml was used, it is designed to separate proteins ranging from 3 to 70 kDa. The Superdex200 16/60 column is designed to separate proteins ranging from 10 to 600 kDa. At first the column was equilibrated with two column volumes of gelfiltration buffer. The protein to be separated was concentrated to a volume of 2 ml using an ultracentrifugation device. The sample was loaded in a 2 ml loop and then injected onto the column. The sample was then eluted using a flow of 0.5 ml/min and was collected in 1.5 ml fractions. The fractions containing the desired protein were collected and concentrated using an ultrafiltration device. The same method was used to purify atEDS1-nanobody complexes. AtEDS1 and the desired nanobody were mixed in a ratio of 1:1.5 and incubated on ice for thirty minutes before application on the gelfiltration column. The residual nanobody, separated from the atEDS1-nanobody complex was collected and stored at 4 °C while the atEDS1-nanobody complex was used for crystallisation.

2.2.3. Protein analytical methods

2.2.3.1. Denaturing SDS PAGE

The sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) as described by Laemmli and colleagues 1970, is a method to separate proteins according to their molecular weight. To denature the proteins the sample is treated with a buffer containing the anionic detergent sodium-dodecyl sulphate (SDS) and a reducing agent (e.g., DTT) for 5 minutes at 95 °C. SDS is used to break apart non-covalent bonds and DTT is used to reduce the cysteine-bridges of the proteins. SDS also binds unspecific to proteins in a ratio of 1.4 g SDS per 1 g of protein and therefore provides a strong negative charge for the proteins. To separate the proteins a polyacrylamide matrix crosslinked with the use of bis-acrylamide was prepared. To guarantee an optimal result, a discontinuous gel matrix was used. A stack gel with a pH value of 6.8 on top of the separation gel with a pH value of 8.8, leads to the stacking of the sample at the border of stack and separation gel, this effect leads to sharper bands and a better result. The concentration of acrylamide added to the gel depends on the size of the proteins to be separated, smaller proteins require a higher amount of acrylamide. Ammoniumperoxodisulfate was used as a radical starter with tetramethylendiamid (TEMED) as a catalyst to generate the sulphate-radicals that start the polymerisation. The gels were stored, wrapped in moist paper towels, in the fridge until use. Electrophoresis was performed in running buffer at 45 mA per gel. To determine the size of the tested proteins, a standard with known proteins (PreStained Plus, ThermoFisher) was used. For all steps the Mini-Proteas Tetra system of BioRad was used.

2.2.3.1. Staining of proteins in SDS gels

After the electrophoresis, the protein bands were fixated and stained using a solution containing ethanol, acetic acid and Coomassie-brilliant blue R250. In detail, the gel was immediately after electrophoresis submerged in the Coomassie staining solution and incubated at room temperature for 30 minutes. Afterwards the staining solution was recovered for reuse and the gel was washed using desalted water to remove excess dye. Afterwards the gel was destained by shaking it with destaining solution overnight. The gel was washed with water to remove excess destaining solution and documented using a imaging system (Chemidoc, Biorad).

2.2.3.2. Western blotting

Alternatively, instead of staining all proteins unspecific, a single protein was labelled using antibodies. For that, the protein was first transferred to a polyvinylidene fluoride (PVDF) membrane using electrophoresis. Directly after the electrophoresis a PVFD membrane, that was activated by submerging it in methanol for five minutes, was layered on the gel without creating air bubbles. Gel, membrane, and a pair of filter paper were then fixed in the Mini TransBlot module (BioRad) according to the manufacturer's manual.

The transfer was performed at 100 V for one hour at room temperature using cold Towbin-buffer. Afterwards the membrane was removed, and empty binding sites were blocked by incubating it with 5% skimmed milk powder (w/v) dissolved in TBS-T for one hour at room temperature or at 4 °C overnight. The membrane was then washed briefly two times with TBS-T to remove residual milk before the first antibody treatment. The membrane was then incubated with the primary antibody that binds to the His-tag of the protein for one hour at room temperature or overnight at 4°C on a shaker. The membrane was washed three times for 10 minutes on a shaker with TBS-T, before it was incubated with the secondary antibody that binds the primary antibody as described above. The secondary antibody was linked to a horse-radish peroxidase (HRP) that was used to detect the presence of the antibody. After the secondary antibody treatment, the membrane was once again washed three times for 10 minutes with TBS-T before it was dried between two filter papers. The membrane was covered in freshly prepared enhanced chemiluminescence lightning (ECL) agent and incubated for 1 minute. Excess ECL was removed, and the luminescence was detected using a documentation system until a clear signal was gained.

2.2.3.3. Quantification of protein solutions

The concentration of protein solutions was determined using their absorption at 280 nm caused by the aromatic amino acids, especially tryptophane. This is a quick and easy method to determine protein concentrations that, by using a nanodrop, requires very little material. To gain an exact estimate of the concentration, the extinction coefficient of each protein was determined using the ProtParam online tool (Wilkins et al. 1999). The protein sequence including all tags and linkers was used to determine the coefficient. The protein concentration was then automatically calculated using Lambert-Beers law.

After blanking the photometer with 2 µl of buffer, 2 µl of the protein sample were measured after carefully mixing the protein solution in the vial by pipetting up and down. The measurement was performed three times and the average was used as the final concentration.

2.2.3.4. Analytical gelfiltration.

To estimate the oligomeric state and molecular weight of a protein, analytical gelfiltration was performed using a Superdex200 10/300GL column. After the column was equilibrated with fresh buffer for three column volumes, a sample of the protein of maximal 500 µl was injected on the column and eluted while collecting the eluate in 1.5 ml aliquots. To calibrate the column and gather an estimate of the molecular weight of the analysed sample, proteins of known size and near global tertiary structure were analysed in a separate run. The exact same tubing and sample-loop was used to guarantee an exact calibration. The elution volume was then plotted against the molecular weight of the protein. A linear fit was applied using Microsoft excel and the resulting function was then used to determine the molecular weight of the protein sample.

The method does not work exact for non-globular proteins, as beside the size, also the tertiary structure of the protein influences the elution volume of proteins. To calibrate the column lysozyme (14.3 kDa), bovine serum albumin (BSA, 66 kDa monomer and 132 kDa dimer), lactate dehydrogenase (144 kDa) and ferritin (440 kDa) were used as standard proteins.

2.2.3.5. Thermal shift assay (thermofluor)

The influence of different buffers, additives and salts on the thermostability of proteins can be determined easily by the thermal shift assay. It is based on the shift in fluorescence when a dye like Sypro-orange binds to hydrophobic amino acids of the protein. These residues are accessible to the dye when the protein is unfolded, and the hydrophobic core is exposed. By gradually heating the sample while measuring the fluorescence of the dye a melting curve of the protein can be recorded, and the melting point can be determined. For easier deduction, the derivative function of the melting curve can be used as the melting temperature correlates with the turning point of the original plot. The protocol used is based on the publication by Boivin and colleagues (2013) . Highly purified protein in gelfiltration buffer was used at a concentration of 25 µM or more. 21 µl per well of the cooled buffer to be tested were pipetted in a white real time-PCR plate (Brand) that was stored on ice. Afterwards 2 µl of the protein solution was added to each well. 3 µl of a 5000x stock solution of Sypro orange (Thermo fisher) in dimethyl sulfoxide was added to 237 µl water and mixed thoroughly before 2 µl of the dye-solution was added to each well. The plate was then sealed using a Microseal B Adhesive seal (Bio-Rad). To guarantee that all components were mixed the plate was centrifuged for a few seconds at 500xg. The melting curve was recorded using a CFX96 -real time PCR machine (BioRad). A temperature gradient reaching from 4 °C to 95 °C was applied with a temperature increase of 0.5°C per thirty seconds. The fluorescence was measured before the next heating step. The results were analysed using the CFX Manager™ software (BioRad). The thermal shift assay cannot just be used to check for optimized buffers but also can reveal small molecules binding to the protein in question. More importantly is that the thermostability shown by thermal shift assay correlates with the proteins tendency to form crystals (Vedadi et al. 2006, Ericsson et al. 2006). In a first thermal shift assay in this work several buffers and pH conditions were tested to optimize protein stability. In a second screening different salts and additives were tested to further increase stability and check for potential ligands. Detailed concentrations and buffers can be found in chapter 3.2.6.

2.2.3.6. Lipase activity assay

A quick and easy method to test proteins for a potential lipase-like activity as well as gather kinetic data is the use of chromogenic artificial substrates such as para-nitrophenol esters. As these substrates are cleaved the released para-nitrophenolate can be detected due to its absorption maximum at 412 nm. Purified protein in gelfiltration buffer was adjusted to a concentration of 1 mg/ml. Gelfiltrationbuffer supplied with 0.5 % Triton X-100 was preincubated at 30 °C before 200 µl was added to each well of a 96 well assay plate. Immediately 25 µl of the protein-solution as well as 25 µl of the para-nitrophenol ester dissolved in dimethylformamide at 10 µM were added. The plate was inserted in a plate reader and mixed for 10 second at medium intensity before the absorption at 410 nm was measured every thirty seconds for 60 minutes at 30°C. Lysozyme was used as a negative control, while a lipase from *Mucor miehei* (Sigma Aldrich) was used as a positive control.

2.2.3. Structural biological methods

2.2.3.1. X-ray crystallography

The three most important methods to determine protein structures are nuclear magnetic resonance spectroscopy (NMR) , X-ray crystallography and cryo-EM. And while in the last years an exponentially growing number of structures solved by cryo-EM are published [9] and the quality of the structures determined improved dramatically the last five years [10] ,the majority of protein structures are still solved via x-ray crystallography. The RCSB protein data bank (PDB) does contain as of September 2021 more than 150.000 structures solved by x-ray diffraction while the EMDB (EMBL-EBI) counts for roughly 16.000 datasets.

One of the limitations of cryo-EM is the dependency of a certain minimal size of the proteins analysed. While several structures of protein with a molecular mass below 100 kDa were published (Herzik Jr. et al. 2019) structures of proteins smaller than 200 kDa are rare. X-ray crystallography does not have a limit of molecule size. As the average length of a human protein is 375 amino acids (Brocchieri & Karlin 2005) which equals roughly 45 kDa (average of 120 Dalton per residue), this means that most proteins of a human cell are not feasible for structure solution with cryo-EM. Of course, this does not consider protein complexes.

9 https://www.ebi.ac.uk/pdbe/emdb/statistics_main.html/

10 https://www.nature.com/articles/d41586-020-00341-9

2.2.3.1. Crystal screening

The bottleneck and most challenging part of the process is often the generation of protein crystals. While there are some instances where proteins forms crystals on their own in the cells (Schönherr et al. 2018) most proteins do not form crystals naturally. It is even proposed that difficulties to crystallize proteins are the result of evolutionary pressure (Doye et al. 2004). To gain protein crystals the first step is to produce a highly concentrated, pure sample of monodisperse protein. In the next step an oversaturation of the protein while keeping its structure needs to be achieved. This may lead to nucleation and finally to crystal growth. In this work, to achieve oversaturation the vapour diffusion method, in a sitting drop setting was used. The protein solution was mixed with the precipitant and placed on a bridge surrounded by the precipitant solution in a large reservoir. The whole system was then sealed using tape. As the solution in the reservoir is higher concentrated than the protein-precipitant mixture, water diffuses from the protein mixture to the reservoir leading to an increase in both protein and precipitant concentration. Figure 4 shows both the used sitting drop method as well as its alternative the hanging drop method schematically.

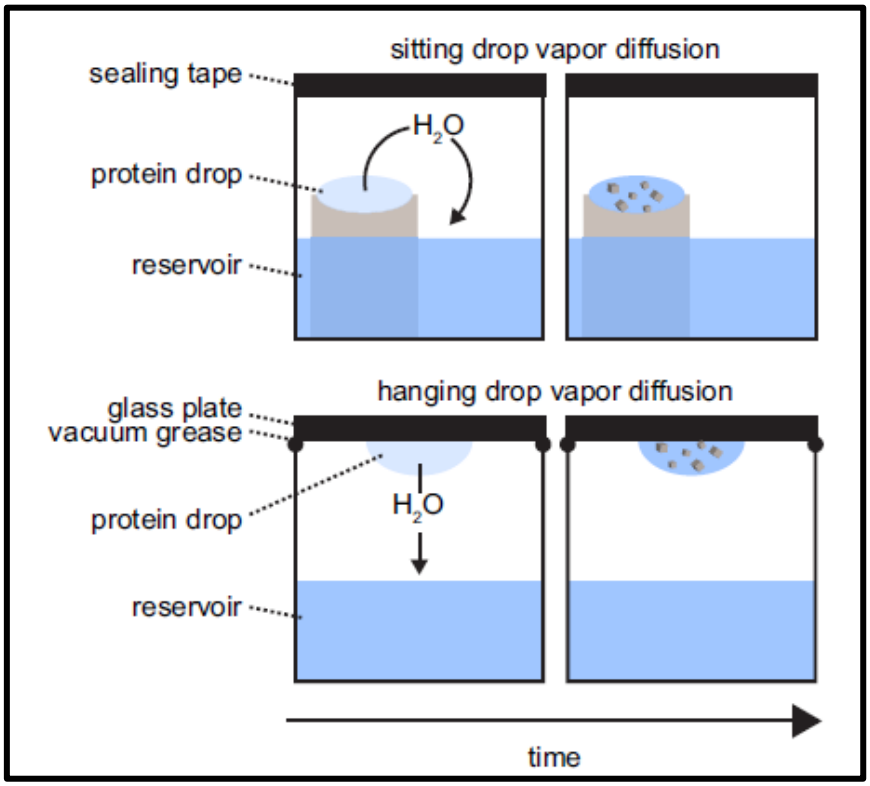

Figure 4: Schematically representation of the hanging drop- and sitting drop- vapor diffusion method for protein crystallization. Protein solution is mixed with the precipitant and placed either on a bridge (sitting drop) or hanged from a seal (hanging drop) surrounded by a reservoir of the precipitant. The system is sealed, and over time water diffuses from the drop in the reservoir, thereby slowly increasing the precipitant and protein concentration in the drop. Taken from the PhD thesis of Stephan Wagner (2013).

The diffusion of water from the drop in the reservoir, leads to a supersaturated solution, nucleation, and crystal growth. This process can be visualised in a phase diagram as shown in figure 5.

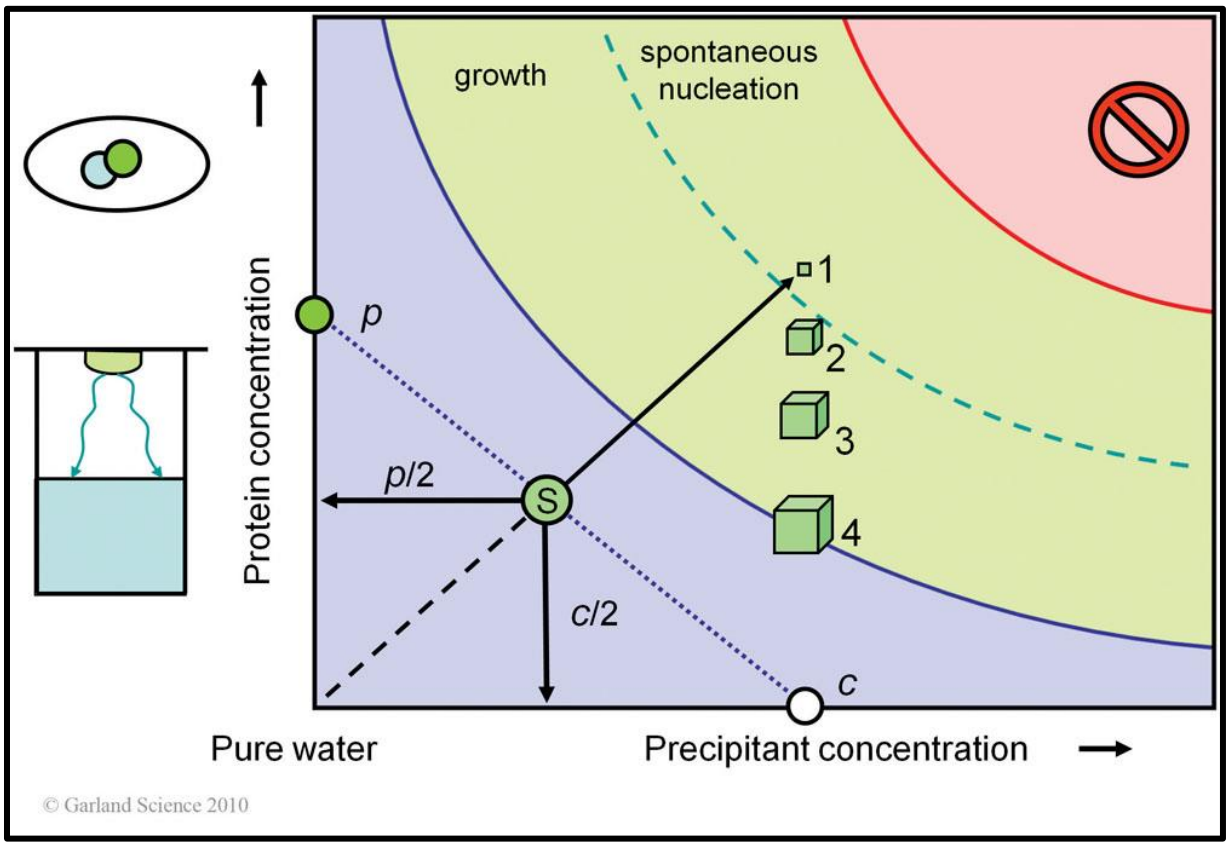

Figure 5:Phase diagram of a successful vapor diffusion crystallization experiment. Protein and precipitant are mixed in equal amounts leading to a undersaturated solution (point S). Over time water diffuses in the reservoir, thereby increasing both protein and precipitant concentration, resulting in spontaneous nucleation (1). The crystals then grow, thereby decreasing the protein concentration until the border to a undersaturated solution is reached (2-4). Taken from Rupp-Biomolecular crystallography. Reproduced with permission from Biomolecular Crystallography by Bernhard Rupp, © 2009-2014 Garland Science/Taylor & Francis LLC

The precipitant solution in which crystallization may occur is different for each protein and cannot be predicted accurately, meaning that a lot of potential conditions need to be tested to succeed. Additionally, several factors including, protein concentration, protein to precipitant ratio, temperature, drop size and the buffer condition each influence the crystallization process. To cover a broad range of conditions, and to approach the crystallization in a systematic approach, commercial screens that contain different precipitant solutions were used in the crystallization experiments. To reduce the amount of protein needed and to test the maximal possible number of conditions a pipetting robot capable of reliable dispersion of drops of about 50 nl was used to mix protein and the crystallization cocktail. The drop size used in this thesis was set to 300 nl with protein to precipitant ratios of 1:1, 2:1 and 1:2. After pipetting the plates were sealed with a clear tape immediately to prevent evaporation. The plates were then stored at 4 °C for several weeks. All experiments were inspected under a binocular regularly and crystals mounted shortly after they stopped growing.

To improve the diffraction quality and size of the crystals, after screening successful conditions were optimized and upscaled (2-3 µl). The conditions of the initial hit were varied by pH-value, salt concentration and precipitant concentration. The plates were stored at 4°C after sealing and incubated for several weeks.

2.2.3.2. Cryoprotection

To minimize radiation damage diffraction measurements are often performed at cryogenic temperatures of about 100 °K. The formation of ice crystals during the freezing process can destroy protein crystals or greatly diminish the diffraction quality (Hope 1988), additionally crystalline ice forms distinct rings on the frames that lead to loss of information and can hinder the analysis of the data. To prevent forming of ice, the crystals were pre-treated with a cryoprotectant solution before flash-freezing in liquid nitrogen. Although often the crystallization cocktail itself might be sufficient to prevent formation of ice crystals. High concentrations of salts, alcohols (e.g. glycerol, 2-propanole, 2,3,-butanediol), sugars (e.g. glucose, sucrose) or glycols (e.g. ethylene glycol, polyethylene glycols of different lengths) can lower the melting point enough to prevent ice formation during flash-freezing. Additionally, mineral oil can be used as cryoprotectant. The crystals can either be transferred directly in the cryoprotectant, or to avoid osmotic shock, transferred in solutions with increasing concentration of the cryoprotectant. In this work, for crystals not grown in conditions that can act as cryoprotection itself, 20 % of either ethylene glycol or glycerol were added to the crystallization conditions. The crystals were transferred into a small drop of the cryoprotectant using a CryoLoop™ (Hampton Research), incubated for roughly thirty seconds and afterwards flash frozen using liquid nitrogen.

2.2.3.3. Mounting of crystals

To perform diffraction measurements crystals were mounted on a nylon loop (Hampton Research) with diameters between 0,05 and 0,3 mm. The protein crystals were then immediately shock frozen in liquid nitrogen and stored in a CrystalCap™ (Hampton Research). The crystals were stored in liquid nitrogen until the diffraction data were collected.

2.2.3.4. Collection of diffraction data

To achieve a near atomic resolution, light with a wavelength of roughly 1 Å, meaning X-rays must be used. When X-rays hit the crystal, they are scattered. The scattered rays then will interfere with each other, this interference can either be destructive or constructive. While destructive interference leads to cancelling out of the rays, constructive interference leads to an increased amplitude of the resulting ray as shown in figure 6.

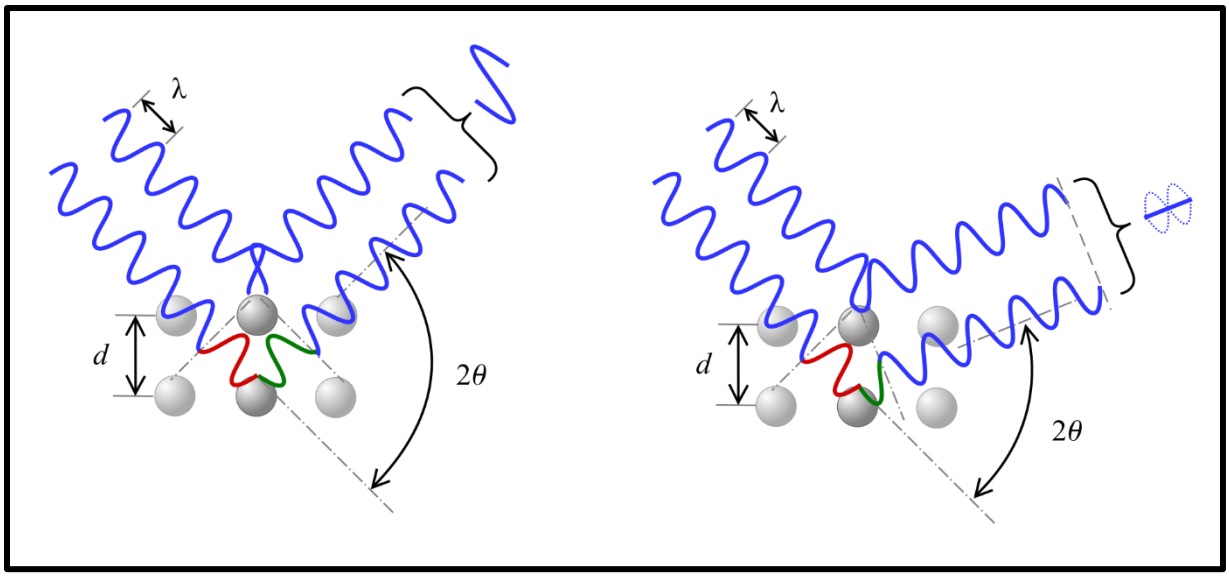

Figure 6: Graphical representation of Bragg`s law. On the left constructive interreference occurs, resulting in a heightened amplitude of the rays. On the right destructive interference cancels out the rays. Taken from Wikipedia [11]

Constructive interference occurs when Braggs law is fulfilled. Braggs law can be expressed as:

$$2dsin\theta = n\lambda$$

With λ the wavelength, d the distance between two lattice points and n an integer number. This means that whenever the difference in the path length, for successive lattice points (2dsinθ) equals an integer number of the wavelength λ, constructive interference occurs. Large angles θ therefore contains the most information about high resolutions.

Diffraction data were collected at different synchrotrons. A synchrotron is in short, an electron accelerator that uses the bremsstrahlung that occurs as the electrons are forced to change direction in a magnetic field to produce electromagnetic radiation. The desired wavelength is filtered using a monochromator and is focused in a small intense beam for use on the crystal. The crystal was mounted on a goniometer and cooled using a nitrogen stream of 100 °K. For tuneable beamlines with variable wavelengths 0,976 Å was chosen. The degree the crystal was turned per frame as well as the number of frames was dependent on the symmetry of the crystal. The strength of the transmission was dependent on the beamline as well as the quality of the crystal. To test the diffraction quality of the crystals, at first four test shots were taken, turning the crystal 90 ° between the shots. This also checked that the crystal was correctly placed in the beam and diffracts equally in all orientations.

11 https://en.wikipedia.org/wiki/Bragg%27s_law

The software at the beamline then usually calculated a recommended recording strategy including exposure time, number of frames, degree of turning per frame and the strength of the beam-transmission. The recommendation was usually followed with slight corrections.

2.2.3.5. Data processing, model building and refinement.

Due to the lack of refractive lenses for x-rays, data are recorded in reciprocal space. As the aim of crystallographic experiments is to gather an electron density map in real space, a Fourier transformation must be performed on the recorded reflections. Unfortunately, information on the structure factor phases is lost during the experiments and cannot be recorded directly. Therefore, information on the phases must be determined differently. The phases can be obtained experimentally using heavy metal derivates of the crystals in isomorphous replacement or using anomalous dispersion. Alternatively, already determined phases from highly similar proteins can be used in the molecular replacement method. In this thesis, phases experimentally determined previously (Wagner et al. 2013) were used for the atEDS1 part of the nanobody complexes as well as the N-terminal vvEDS1-structure. Phases for the nanobody were gathered from a structure published in the PDB (4KRN).

The images recorded at the synchrotron were processed using the AUTOPROC pipeline (Vonrhein et al. 2011), that includes XDS (Kabsch, 2010) for indexing and integration, POINTLESS (Evans, 2006) as well as AIMLESS (Evans & Murshudov 2013) for space group determination and scaling as well as STARANISO (Tickle et al.2018) to analyse and compensate for anisotropy. Molecular replacement was done using PHASER (McCoy et al. 2007) as part of the PHENIX suite (Liebscher et al. 2019). The atEDS1 part of the atEDS1-atSAG101 structure (PDB 4NFU) as well as the nanobody structure 4KRN were used as search model for the atEDS1-nanobody structures. For the vvEDS1-N-terminus structure the N-terminal domain of the atEDS1-structure of the atEDS1-atSAG101 complex was used as search model. The vvEDS1-N-terminus structure was then build using AutoBuild (Terwilliger et al. 2008). All structures were then further optimized by several rounds of PHENIX refinements (Liebscher et al. 2019) followed by manual model building using COOT (Emsley et al. 2010). The atEDS1-nanobody structures were deposited in the PDB with the codes 6I8G (EDS1-1AT72), 6Q6Z (EDS1-1AT21) and 6I8H (EDS1-1AT15).

3.Results

3.1. Structure determination of atEDS1-nanobody complexes

3.1.1 Expression and purification of nanobodies

Single chain antibody fragments (nanobodies) from *Llama glama* were generated by Christine Tölzer (Institute of biochemistry, University of Cologne, AG Prof. Niefind) in cooperation with VIB (Brussel, Belgium) previously (Voss et al. 2019). For that, purified His-tagged atEDS1 (EDS1 from *A.thaliana*) was injected in a llama six times over a 35-day period. A cDNA library was generated from the lymphocytes of the llama using reverse transcription. The nanobody sequences were isolated, amplified via PCR and inserted into a suitable vector. The EDS1 specific nanobodies were then isolated via three rounds of phage display. The resulting EDS1 specific nanobodies were sequenced and the sequences analysed and grouped. The specificity of the nanobody and therefore the epitope is determined by three loops with a high degree of variability between the different groups. Nanobodies of the same group are identical in the epitope binding site but show slight differences in the backbone of the nanobody. Therefore, nanobodies of the same group bind exactly to the same epitope.

The nanobodies were overexpressed recombinantly in *E. coli* WK6 cells and isolated as described in 2.2.2.3. In a first step, one nanobody of each group was selected randomly, expressed, and purified using at least two litres of culture. While some nanobodies could easily be purified with a yield of 0.5-2 mg/l of culture, others were not expressed as indicated in table 13. If a different nanobody of the same group was available, in a next step this one was expressed and purified and so on. Nevertheless, no member of the groups 3, 6, 7, 18, 20 and 22 was successfully purified from the *E.coli* WK6 strain as shown in table 13.

Table 13: Results of the initial expression of atEDS1 specific nanobodies. Members of the same group bind to the same epitope but have slightly different backbones. Black nanobodies were successfully expressed at first try. Blue nanobodies were not tested as a different member of the group was available. Red nanobodies were not expressed in an initial purification using E. coli WK6 cells.

Group	Member(s)
1	1AT80, 2AT3
2	1AT3
3	1AT23
4	1AT55, 1AT73, 2AT17, 2AT8
5	1AT15, 1AT22, 1AT85, 1AT29
6	2AT2
7	1AT59
8	2AT11, 2AT19, 2AT13, 1AT66, 1AT45, 1AT94, 2AT22
9	1AT5, 1AT67
10	1AT1, 2AT16, 2AT18, 1AT38
11	1AT74
12	1AT13
13	1AT27
14	1AT20, 1AT72, 2AT6, 1AT78, 2AT9
15	1AT61, 1AT84
16	2AT24
17	1AT21, 2AT7, 1AT81
18	1AT86
19	2AT4
20	2AT20
21	1AT32
22	2AT26

One obvious drawback of using *llama glama* nucleotide sequences for protein expression in *E. coli* is the different codon usage in both organisms. This may result in an accumulation of rarely used codons, that might explain the difficulties in expressing certain nanobodies. To overcome this problem WK6 cells were transformed with the pRARE plasmid taken from the Rosetta™ *E.coli* strain, that encodes for several tRNAs rarely used by *E. coli*. Adding this plasmid resulted in the successful expression and purification of 1AT23, 2AT2, 1AT59, 1AT86 and 2AT20, meaning one nanobody of each group, except group number 22 (2AT26) was purified and available for testing. The resulting yield was typically 0,3-1 mg/ litre of LB medium. The protein isolated usually showed a high degree of purity, therefore no second purification step was required (figure 7).

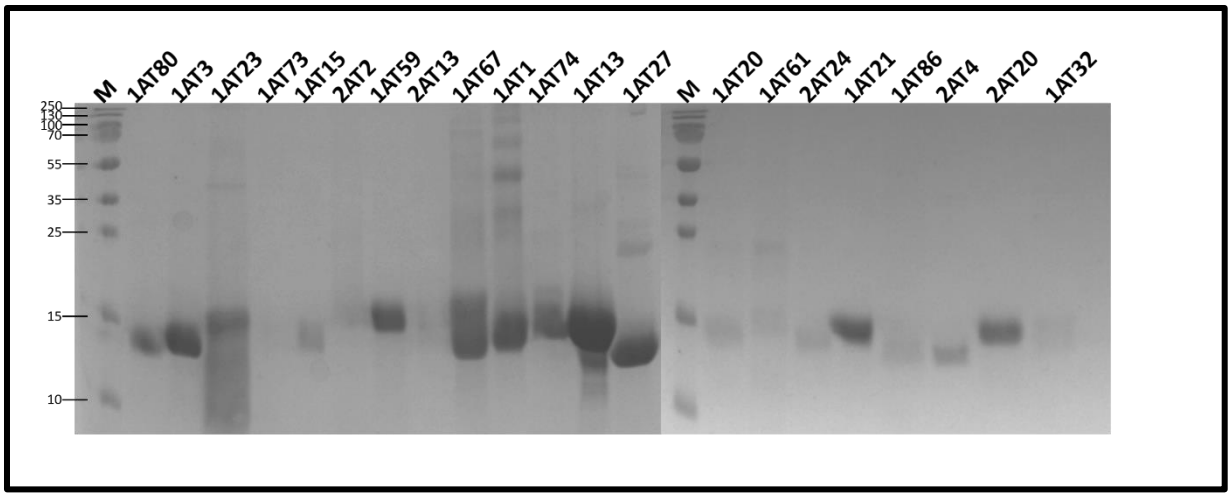

Figure 7: Coomassie stained SDS-PAGE of atEDS1 specific nanobodies. Nanobodies from llama were expressed in E.coli and purified via NiNTA chromatography. The proteins were applied according to the group-numbers as determined by the manufacturer and shown in table 13.

3.1.2. Selection of nanobodies for crystallization

The nanobodies were prepared and purified with the aim to act as a crystallization chaperone for atEDS1, as crystals of the protein itself have poor and insufficient diffraction qualities (Wagner 2013). As using nanobodies randomly for crystallisation trials seemed to be inefficient, experiments were performed to determine promising candidates for crystallization. The thermal shift assay is a quick method to determine the thermal stability of a protein or a protein complex. Since more stable proteins tend to crystallize better (Dupeux et al. 2011), a subset of EDS1-nanobody complexes were investigated for thermal stability. Figure 8 shows the melting curves (Figure 8A) and derivative results (Figure 8B) of these assays. While pure atEDS1 shows a single melting point at around 49 °C, all eleven analysed nanobody-EDS1 complexes show two distinct melting points.

The first melting point was determined to be 39 °C, with the atEDS1-1AT15 complex as exception at 43 °C. A second range of melting points was detected between 58 °C (1AT3) and 67 °C (1AT13). At this point no selection based on the melting curves was done, as the reason for the two distinct melting points is unknown.

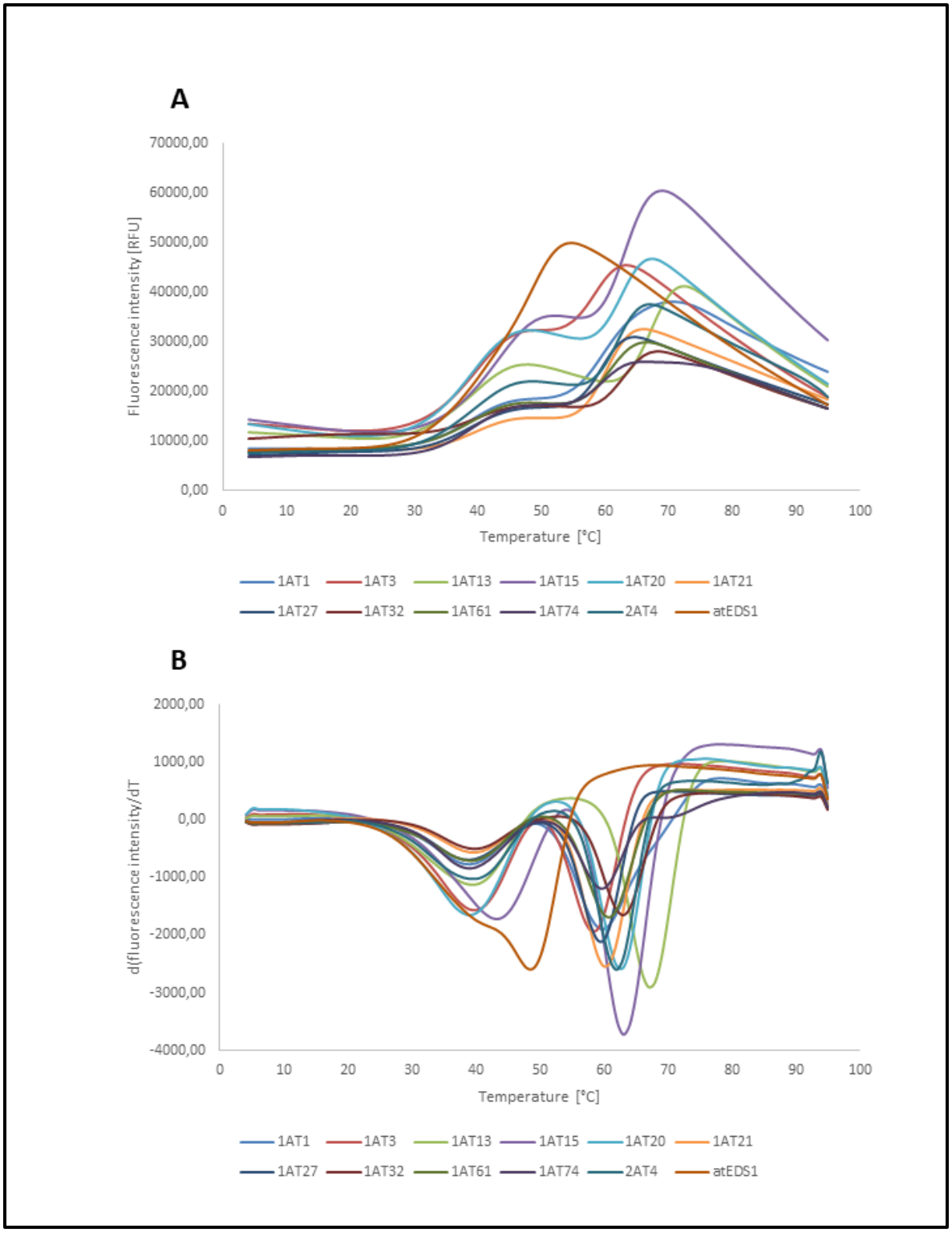

Figure 8: Thermal shift analysis of atEDS1-nanobody complexes by differential scanning fluorometry. A) Measured fluorescence plotted directly against the temperature. B) First derivatives of the direct curves in panel A for easier analysis. Equimolar amounts of atEDS1 and different nanobodies were mixed and incubated on ice. Afterwards SYPRO-orange was added, and the plate was heated stepwise while measuring the fluorescence in a real-time PCR machine. The brown line is the negative control with just atEDS1. The other lanes represent the atEDS1-nanobody complexes.

In a next selection step, the binding site of the nanobody was mapped. The N-terminal domain of atEDS1 (1-384) can be expressed separately in *E.coli* using the same protocol as the full length atEDS1. The C-terminal domain is insoluble when expressed separately in *E.coli* and so a purification is not possible (Wagner 2013). To determine if the nanobodies bind to the N-terminal or the C-terminal domain, size exclusion chromatography experiments were performed. An excess of N-terminal atEDS1 was mixed with the nanobody and a gelfiltration was performed. Nanobodies with an affinity for the N-terminal domain would bind and form a complex, resulting in a single peak. While nanobodies binding to the C-terminal domain or a different protein would be separated on the column, resulting in two distinct peaks. Nanobodies not binding the N-terminal domain were, in a follow up step, mixed with an excess of full-length atEDS1 and again analysed via size-exclusion chromatography to guarantee a binding to atEDS1.

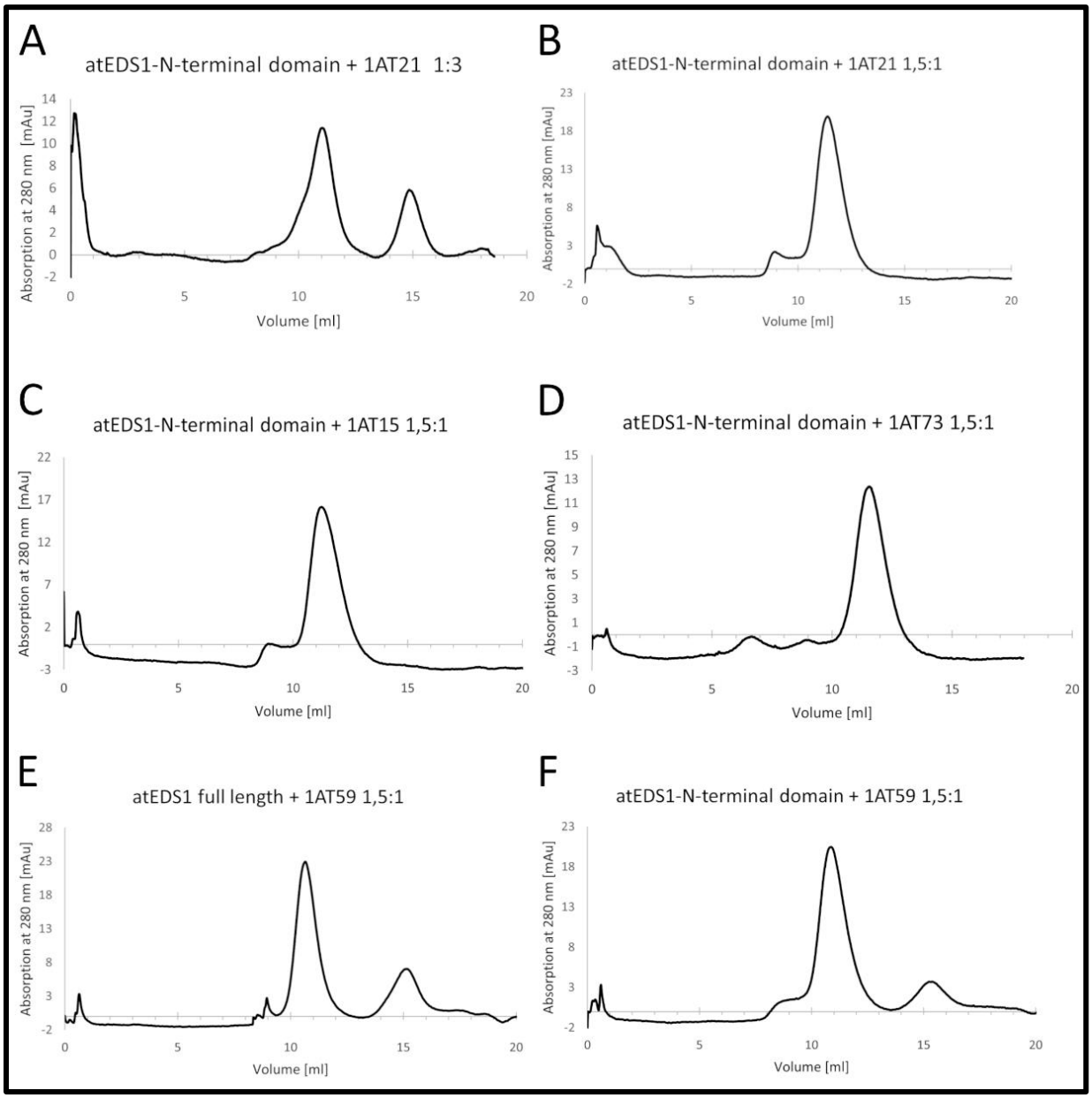

Figure 9: Chromatograms of the gelfiltration to determine the binding site of different nanobodies to atEDS1. Different ratios of either atEDS1 full length of the N-terminal domain of atEDS1 both expressed in E. coli, were mixed, and applied to a Superdex75 size exclusion column. 1AT15, 1AT21 and 1AT73 were found to bind the N-terminal domain of atEDS1 while apparently 1AT59 does not bind to atEDS1 at all.

Figure 9 shows the result for the analysis of the nanobody 1AT21, 1AT15, 1AT73 and 1AT59. Part A proves that the nanobody and the nanobody-EDS1 complex can be separated by the Superdex75 16/60 column, as two distinct peaks at 12 ml and 15 ml appeared, when an excess of nanobody to the N-terminal domain of EDS1 is applied. Part B shows just one distinct peak for the atEDS1 N-terminal domain-nanobody complex as atEDS1 is mixed in excess with the nanobody 1AT21. As no free nanobody is detected, this indicates that 1AT21 indeed binds to the N-terminal domain of atEDS1. Parts C and D of figure 9 show the same result for 1AT15 and 1AT73, indicating that these two nanobodies also bind the N-terminal domain of atEDS1. It was found that all nanobodies except 1AT59, 1AT13 and 2AT24 bind to the N-terminal domain of atEDS1. The three nanobodies were then tested as described above with the full length atEDS1. This resulted again in two distinct peaks as shown for 1AT59 in figures 9E and 9F. This indicates that the nanobodies did not bind atEDS1 at all, or not strong enough to build a complex stable enough to last on the gelfiltration.

Notably both the N-terminal domain, as well as the full length atEDS1 were found at similar elution volumes. Although the column was not calibrated for these experiments this points towards a dimerization of the N-terminal domain of atEDS1. This puzzling result can be explained taking in consideration the results involving vvEDS1 (EDS1 from *V.vinifera*, referenz) that also show a strong dimerization for the N-terminal domain. At this point it was sufficient to know that the nanobodies do not bind to the C-terminal domain of atEDS1.

As no nanobody binding to the EP-domain was identified and no effect on the thermal stability of the complex was found with the thermal shift assays, random crystallization trials were performed. As Christine Tölzer was, in a first experiment, able to grow small crystals of an atEDS1-1AT15 complex and this nanobody had a visible effect on the thermal stability (Figure 8) 1AT15 was chosen as a first candidate for crystallization.

3.1.3. Structure determination of the atEDS1-1AT15 complex

The atEDS1/1AT15 complex was purified as described in 2.2.2.5. and concentrated to a final concentration of 4.1 mg/ml. The protein was then applied to several crystallization screens at room temperature and 4 °C. After three days, crystals appeared in roughly fifteen conditions in different size and quantity. All crystals appeared in a hexagonal shape and varied from roughly 0,05 to 0,15 µM as shown in figure 10.

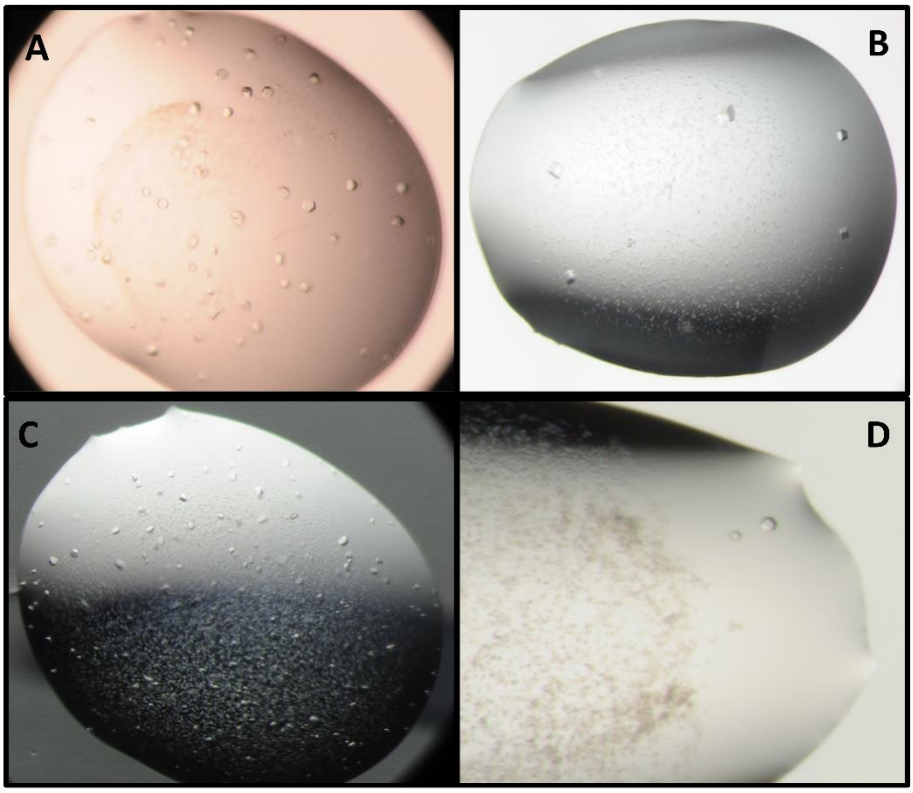

Figure 10: Representative pictures of atEDS1-1AT15 crystals. Crystals appeared after three to five days incubation. Crystals were grown in sitting drop crystal screening plates with a drop size of 300 nl at different protein/precipitant ratios. Photos were taken using a camera mounted on a binocular.

The crystals were observed for ten days before cryoprotection, mounting and storage in liquid nitrogen, as no further growth was observed after seven days. Several crystals were measured at the ESRF (BM14, Grenoble, France) and the EMBL (P13, Hamburg, Germany). Two datasets were recorded at 6.1 Å and 3.7 Å resolution, a vast improvement to all former attempts to crystallize atEDS1 on its own. These attempts resulted, despite much effort by Stephan Wagner (Wagner 2013) and Christine Toelzer (unpublished results), in poor diffracting crystals with resolutions around 10 Å.

The best crystal of the complex was grown in a condition containing 12.5 %(v/v) 2-methyl-2,4,-pentadiol (MPD), 12.5 %(v/v) PEG1000, 12.5 %(w/v) PEG3350, 30 mM $MgCl_2$, 30 mM $CaCl_2$, 61.2 mM 2-(N-morpholino)-ethane sulfonic acid (MES), 38.8 mM imidazole, pH 6.5. In a next step it was attempted to gain larger and better diffracting crystals by varying the concentrations of the precipitants and salts as well as enlarging the drop size. Despite using a wide variety of conditions including different drop size, protein to precipitant ratios as well as different protein batches no crystals were observed in all attempts. To enhance crystal formation in a next attempt additive screens were used in combination to the successful screening conditions mentioned above. This also did not lead to crystal formation.

Another attempt to gain better diffracting crystals was the use of micro seeding with several of the initially successful conditions. This once again was unsuccessful. Finally, the initial screening was repeated two times with different protein batches, once again without crystals forming.

Due to the failure of all crystallization optimization attempts, the determination of the atEDS1/1AT15 complex structure proceeded with the aforementioned 3.7 Å diffraction data set that was processed and analysed as described previously (see 2.2.3.5.). The phase problem was solved using molecular replacement using a nanobody structure (PDB: 4KRN) as well as the atEDS1 part of the atEDS1/atSAG101 structure (PDB 4NFU). This was followed by several rounds of manual model building and refinement. The final structure was published (Voss et al. 2019) and deposited in the PDB with the identifier 6I8H. The statistics are shown in table 14.

Table 14: Statistics of the atEDS1-1AT15 structure, refinement, and data collection.

atEDS1/1AT15		
PDB code		6I8H
X-ray diffraction data collection		
Wavelength [Å]		0,97625
Synchrotron and beamline		EMBL/PETRA-III P13
Space group		$P6_322$
Unit cell	a, b, c [Å]	145.887, 145.887, 152.770
	α, β, γ [°]	90.0, 90.0, 120.0
EDS1-1AT15 dimers per asymmetric unit		1
Resolution [Å] (highest resolution shell)		97.361-3.682 (4.050- 3.682)
R_{sym} [%]		11.1 (134.9)
CC ½		0.998 (0.536)
Signal to noise ratio (1/σ)		9.8 (1.5)
Number of unique reflections		8246 (411)
Completeness (spherical) [%]		75.5 (15.6)
Completeness (ellipsoid) [%]		93.0 (59.5)
Multiplicity		6.3 (7.4)
Wilson B-Factor [$Å^2$]		167.66
Structure refinement and validation		
Number of reflections for R_{work}/R_{free}		8237/800
R_{work}/R_{free} [%]		24.2/28.7
Number of protein-atoms		5893
Average B-factor [$Å^2$]		199.31
RMS deviations:		
Bond lengths [Å]		0.004
Bond angles [°]		0.95
Ramachandran plot		
Favoured (%)		95.62
Allowed (%)		3.84
Outliers (%)		0.55

The spacegroup was determined as $P6_322$ with a unit cell of 145.887 Å, 145.887 Å and 152.770 Å and angles of 90°, 90° and 120°. According to the Matthews probability one EDS1-1AT15 complex was placed per asymmetric unit. The resolution ranged from 97.361 Å to 3.682 Å, which allowed the correct placement of the nanobody to the atEDS1 molecule, in contrast to the dataset diffracting to 6.1 Å recorded at the BM14 (ESRF) where the binding site of the nanobody could not be confirmed with confidence. The dataset was significantly anisotropic as reflected in the spherical completeness of 75.5% overall and 15.6 % in the outer resolution shell, compared to 93.0 % and 59.5 % for the ellipsoidal completeness. The model building resulted in a R_{work}/R_{free} value of 24.2/28.7 %. All other quality indicators are shown in table 14 and fit the expected values for a structure of comparable resolution.

3.1.4. Results of random crystallization trials

While the atEDS1/1AT15 complex was selected as a crystallization candidate due to the previous success of Christine Tölzer, further crystallization experiments were performed with atEDS1 in complex with random nanobodies. The complexes were prepared and concentrated to a concentration between 1.5 and 4.5 mg/ml. The resulting protein was then added to as many crystal screens as protein was available. Screening of complexes consisting of atEDS1 and 1AT80, 1AT3, 1AT61, 1AT74, 1AT23 and 1AT1 did not yield any crystals. Crystals could be grown from atEDS1-1AT32 and atEDS1-1AT20 complexes but the resulting diffraction of ~20 Å did not warrant further optimization. As atEDS1-1AT21 and atEDS1-1AT73 complexes were successfully crystallized no further complexes were tested for crystallisation. In conclusion, eight groups of atEDS1 binding nanobodies were not tested in this thesis. All mentioned nanobodies were tested just once, with several crystal screens each.

3.1.5. Structure determination of the atEDS1-1AT21 complex

After the successful crystallisation of the atEDS1-1AT15 complex, the next promising hit consisted of the atEDS1-1AT21 complex. The crystals were discovered as a result of random crystallisation trials. The first screening with a concentration of 3.7 mg/ml resulted in roughly twenty successful conditions with crystals of different sizes ranging from 0.05 to 0.2 µm. The crystals appeared after two to five days of incubation at 4 °C. Conditions mostly consisted of Bis-Tris buffered solutions pH 7.5 or 8.5, together with PEG3350 and different salts (sodium sulfate, sodium citrate, sodium nitrate, sodium tatrate). Unfortunately, the diffraction of these crystals was poor with resolutions around 6 Å. The best resolution was derived from crystals from the “Morpheus” screen (Molecular dimensions) that had the additional advantage of not needing additional cryoprotection. A dataset with a resolution of 3.47 Å was gained form such a crystal grown in a condition containing 10 %(w/v) PEG4000, 20 %(v/v) glycerol, 30 mM $MgCl_2$, 30 mM $CaCl_2$, 60.1 mM TRIS, 39.1 mM BICINE, pH 8.5.

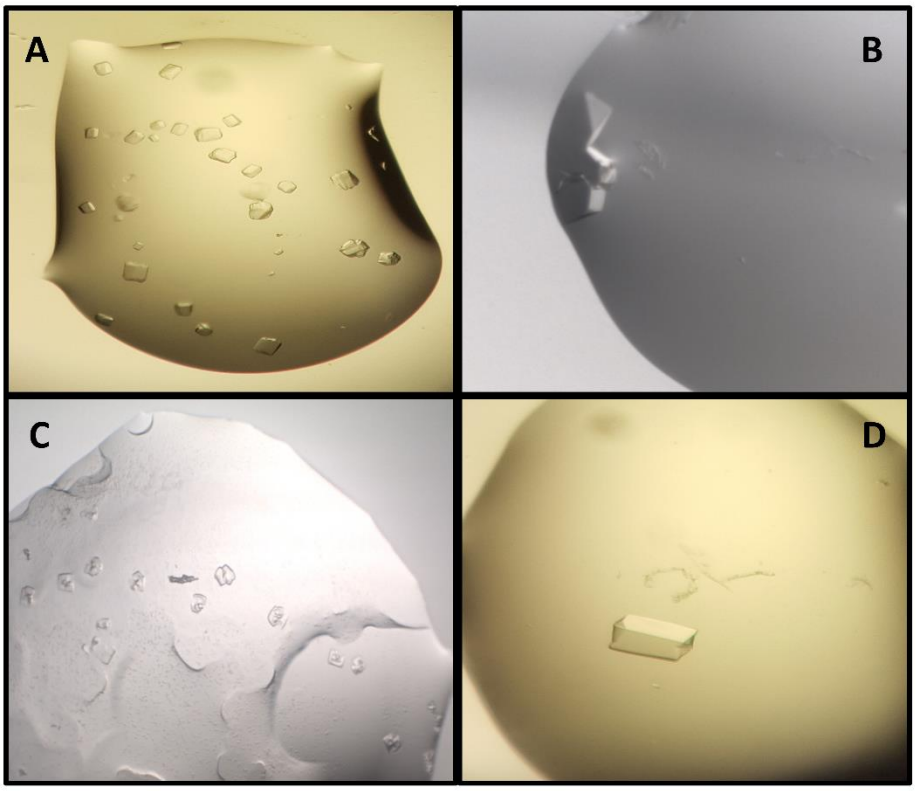

Figure 11: Representative pictures of atEDS1-1AT21 protein crystals. Crystals appeared after two to three days incubation at 4°C. Crystals were grown in sitting drop crystal screening plates with a drop size of 300 nl at different protein/precipitant ratios. Photos were taken using a camera mounted on a binocular.

Once again it was attempted to reproduce the crystals in a larger drop in the hope to grow larger and better diffracting crystals, including several variations of the salt and precipitant solutions as well as the pH values. None of the attempts proved to be fruitful as no crystals were formed. Micro seeding using some intergrown crystals from the initial screening also did not yield crystals. The screening was then repeated using the same procedure as the first attempt. Again, crystals were found in several conditions in varying sizes. Nevertheless, none of these crystals provided a better resolution than 3.7 Å. The dataset was solved as described above for the atEDS1-1AT15 structure. Table 15 shows the statistics of the data collection and refinement.

Table 15: Statistics of the atEDS1-1AT21 structure, refinement and data collection.

atEDS1-1AT21		
PDB code		6Q6Z
X-ray diffraction data collection		
Wavelength [Å]		0,97625
Synchrotron and beamline		ESRF (ID23-1)
Space group		P321
Unit cell	a, b, c [Å]	142.192, 142.192, 97.802
	α, β, γ [°]	90.0, 90.0, 120.0
EDS1-1AT21 dimers per asymmetric unit		1
Resolution [Å] (highest resolution shell)		76.586-3.473 (3.646-3.473)
R_{sym} [%]		18.2 (581.3)
CC ½		0.993 (0.354)
Signal to noise ratio (1/σ)		12.1 (1.3)
Number of unique reflections		12643 (634)
Completeness (spherical) [%]		83.8 (31.7)
Completeness (ellipsoid) [%]		90.8 (57.1)
Multiplicity		19.5 (20.1)
Wilson B-Factor [$Å^2$]		158.88
Structure refinement and validation		
Number of reflections for R_{work}/R_{free}		12589/766
R_{work}/R_{free} [%]		21.9/26.8
Number of protein-atoms		5880
Average B-factor [$Å^2$]		188.2
RMS deviations:		
Bond lengths [Å]		0.004
Bond angles [°]		1.06
Ramachandran plot		
Favored (%)		96.0
Allowed (%)		3.59
Outliers (%)		0.41

The resolution ranged from 76.586 to 3.473 Å allowing the placement of the nanobody in the electron density. The dataset proved to be slightly anisotropic with a spherical completeness of 83.8 % overall and 31.7 % in the highest resolution shell (3.646-3.473 Å) and an ellipsoidal completeness of 90.8% overall and 57.1 % in the highest resolution shell. The R_{work} after the final round of model building and refinement was 21.9 % with a R_{free} value of 26.8 % with all other statistics in the expected range for a structure of comparable resolution, proving the model to be a valid one. The structure was deposited in the PDB with the identifier 6Q6Z.

3.1.6. Structure determination of the atEDS1-1AT73 complex

The next complex to be crystallized successfully consisted of the nanobody 1AT73 bound to atEDS1. The complex was added to the crystallisation screens at a concentration of 2.8 mg/ml, after two to five days crystals appeared in a broad range of conditions. The crystals grew in patches of intergrown, thin plates (Figure 12).

The crystals grew in conditions containing PEG3350, HEPES or Bis-Tris at pH values between 6.5 and 8.5 as well as several salts including sodium sulphate, sodium citrate, sodium tartrate, sodium formate and sodium nitrate. These initial crystals were mostly too small and intergrown to permit easy mounting in a nylon loop and where therefore not used further. In a next step several of the successful conditions were once again reproduced and the concentrations of salts and precipitant varied. Crystal growth was observed on several plates in a variety of conditions, mostly containing lower concentrations of PEG3350 than the 25% (w/v) of the screening conditions. Once again, the crystals grew in patches of intergrown plates, although the size was increased in some conditions compared to the initial screening.

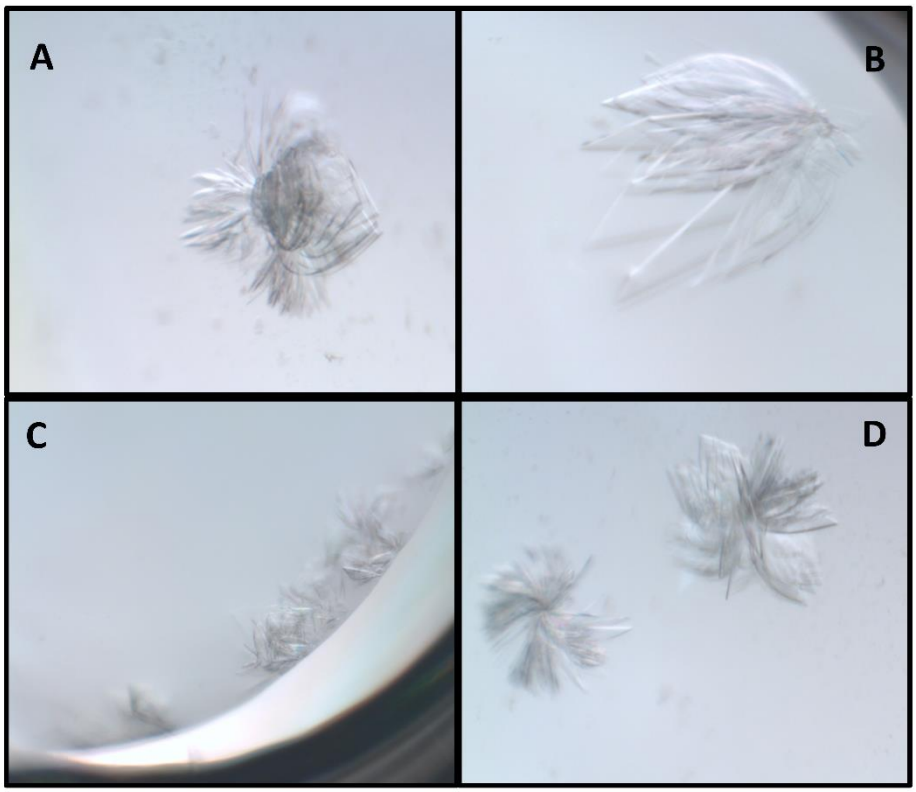

Figure 12: Representative pictures of the atEDS1-1AT73 protein crystals. Crystals appeared after two to three days incubation at 4°C. Crystals were grown in sitting drop crystallization plates with a drop size of 2µl at different protein/precipitant ratios. Photos were taken using a camera mounted on a binocular.

To mount a single crystal the clusters of plates were carefully separated using a needle before mounting a single plate for flash freezing and diffraction measurements. Several datasets were collected with a crystal grown in 17.5% PEG3350, 0.2 M sodium citrate, 0.1 M Bis-Tris, pH 8.5 providing the best results with a resolution of 2.34 Å.

The quality of the dataset allowed to build a much more detailed model with a confident placement of the sidechains of the protein core. Table 16 shows the statistics and results of the data collection, model building and refinement.

Table 16: Statistics of the atEDS1-1AT73 structure, refinement, and data collection.

atEDS1-1AT73		
PDB code		6I8G
X-ray diffraction data collection		
Wavelength [Å]		0,9770
Synchrotron and beamline		ESRF (MASSIF-3)
Space group		C2
Unit cell	a, b, c [Å]	177.504, 68.232, 105.273
	α, β, γ [°]	90.0, 123.44, 90
EDS1/1AT73 dimers per asymmetric unit		1
Resolution [Å] (highest resolution shell)		83.773-2.344 (2.565-2.344)
R_{sym} [%]		11.7 (87.0)
CC ½		0.995 (0.571)
Signal to noise ratio (1/σ)		8.0 (1.5)
Number of unique reflections		31550 (1577)
Completeness (spherical) [%]		71.1 (15.1)
Completeness (ellipsoid) [%]		92.4 (62.2)
Multiplicity		3.6 (3.7)
Wilson B-Factor [$Å^2$]		42.14
Structure refinement and validation		
Number of reflections for R_{work}/R_{free}		30242/1295
R_{work}/R_{free} [%]		19.4/22.2
Number of non-H-atoms		6170
Protein		6001
Water		169
Average B-factor [$Å^2$]		56.06
Protein		56.42
Water		43.44
RMS deviations:		
Bond lengths [Å]		0.002
Bond angles [°]		0.53
Ramachandran plot		
Favoured (%)		95.68
Allowed (%)		3.65
Outliers (%)		0.68

The dataset was strongly anisotropic with a spherical overall completeness of 71.1% versus 92.4 % for the ellipsoidal dataset. The model building resulted in a R_{work} of 19.4% and a R_{free} of 22.2%. Due to the increased resolution compared to the atEDS1-1AT15 and atEDS1-1AT73 structure allowed the placement of 169 coordinated waters in the structure.

3.1.7. Study of atEDS1-aggregation over time.

Solutions of purified atEDS1 tend to form aggregates over time, which results in a white precipitate that occurs roughly 48 hours after the purification, if the protein was stored at 4°C, or much quicker at room temperature. Buffer conditions to improve the stability of atEDS1 in solution had been established prior to my work (Wagner 2013; C. Tölzer, unpublished work). Beside a low salt concentration and a pH value between 7.5-8.5, glycerol as well as reductive agents such as DTT or TECEP tend to stabilize the protein. The tendency to form aggregates might also explain the various reports that found a self-interaction of atEDS1 as described in the introduction. I found that reductive agents were especially important to produce a monodisperse atEDS1 solution without aggregates. As DTT has a low half-lifetime in solution, even the equilibration of gelfiltration columns the night before, led to more aggregates compared to a column equilibrated just before use. To underline the importance of reductive agents for the stability of atEDS1 a series of gelfiltrations was performed as described in 2.2.3.4.

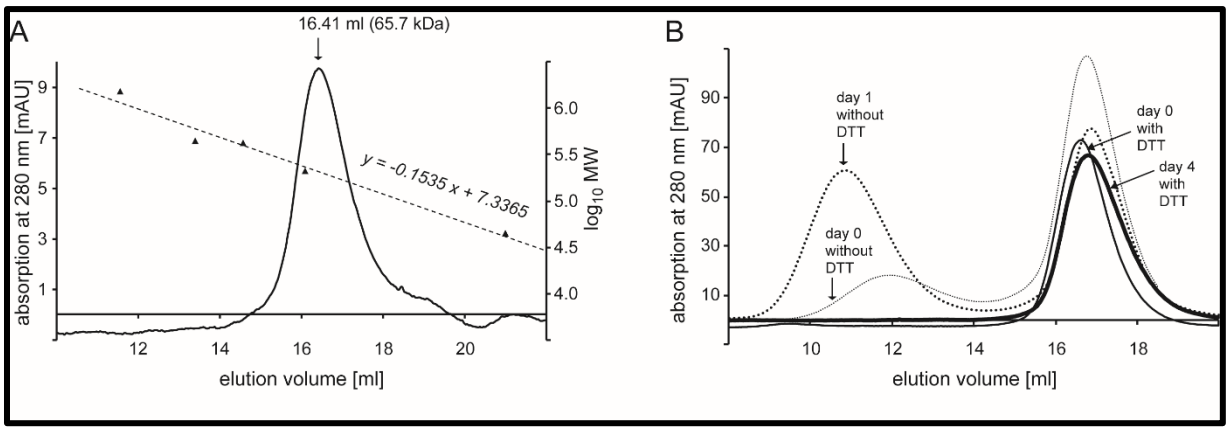

Figure 13: AtEDS1 forms aggregates over time. Figure A shows the chromatogram of a gelfiltration of atEDS1 on a Superdex200 13/30 column that was calibrated using several proteins of known mass. AtEDS1 was found to be roughly 65 kDa large, hinting to a monomeric quaternary structure. Figure B shows the same experiments after zero to four days when either fresh DTT was supplied daily or none DTT was supplied after the initial purification. In the absence of DTT atEDS1 rapidly forms aggregates, after two days the sample blocked the column and could not be further used.

It was found that, if fresh DTT is supplied continuously, even after four days no aggregates were detected in the gelfiltrations as shown in figure 13B. Just a single peak at the expected elution volume of atEDS1 was found. In contrast, if no fresh DTT is supplied after the initial purification, immediately a significant peak close to the void volume of the column is found. After 24 hours roughly half of the protein seemed to be precipitated and forms a peak around the void volume of the column. The sample applied after 48 hours was aggregated to a degree that clogged the column.

Taken together, these results suggest that formation of disulfide-bridges drive the aggregation of atEDS1. To finally prove this point further analysis will have to be performed.

3.2. Expression and characterization of EDS1 and PAD4 orthologues

3.2.1. Expression of an atPAD4 inactive mutant

As described in the introduction, EDS1 binds to its family members SAG101 and PAD4. After the atEDS1-atSAG101 complex was solved by Wagner et al. (2013) and the structure of the atEDS1 monomer in this work, the next logical step is to purify and crystallize an atEDS1-atPAD4 complex. Unfortunately, while atPAD4 (PAD4 from *A.thaliana)* can be easily produced by *E.coli,* we were never able to gain soluble protein, as all protein was always detected in the insoluble fraction of the cell lysate after centrifugation. Previous efforts in the group to increase atPAD4 solubility had comprised refolding from inclusion bodies, expression in *P.pastoris*, several solubility tags (including glutathione-S-transferase, maltose binding protein and thioredoxin), co-expression with atEDS1, different *E.coli* strains as well as co-expression of chaperones, but all this had proved to be unsuccessful. In this work, I supplemented these attempts by the expression of atPAD4-Ser118Ala, a point mutant designed with the following rationale: AtPAD4 just like atEDS1 has an intact catalytic triad that is conserved among species and therefore has the potential to be an active lipase-like enzyme. As such, PAD4 in its natural form might be toxic for the bacteria cells, leading to accumulation of cells that express insoluble atPAD4. To test this hypothesis the catalytic serine Ser118 was mutated to an alanine, thereby disrupting the catalytic triad and all potential lipase-like activity.

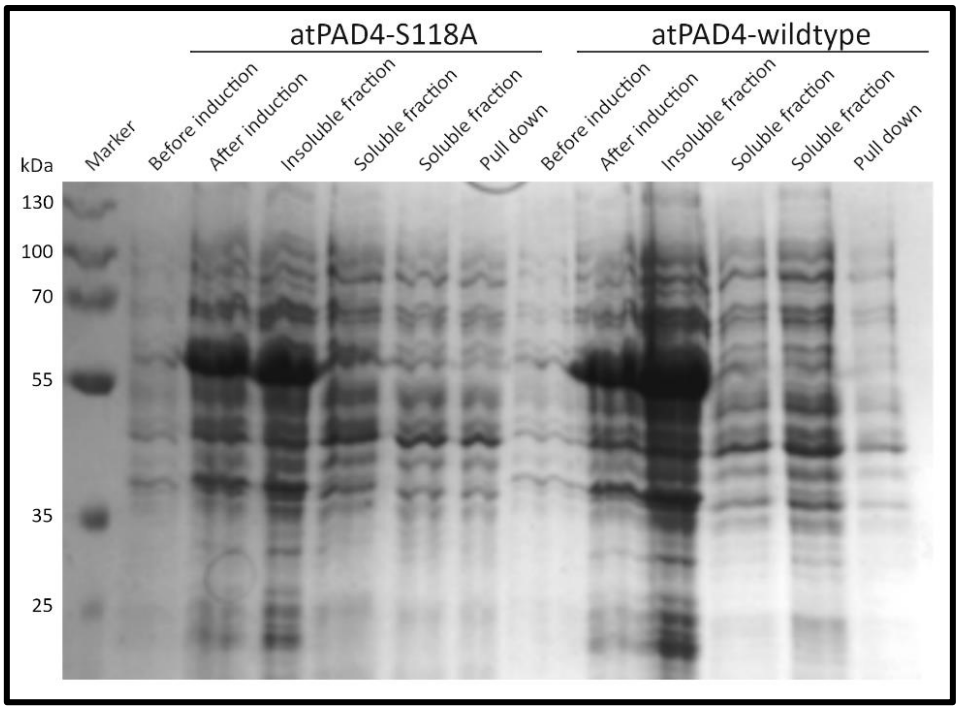

Figure 14: Testexpression of atPAD4-wt and atPAD4-S118A (60 kDa) in E. coli. Both protein variants were expressed in a small-scale culture of BL21 (DE3) cells, samples were taken before and after induction with IPTG and incubation at 18 °C overnight. Cells were then lysed, and insoluble debris removed via centrifugation. A small amount of Ni-NTA resin was incubated with the lysate for 1 h. The supernatant was discarded after centrifugation and the potential bound proteins eluted using SDS-sample buffer. No washing steps were performed. Both protein variants were expressed in high amounts but were solely found in the insoluble fraction after lysis of the cells.

The test expression is shown in figure 14. Both the wildtype and the Ser118Ala mutant protein are overexpressed in the *E.coli* cells after induction with IPTG. Unfortunately, both variants are found solely in the insoluble fraction after cell lysis and centrifugation, indicating that the mutation had no influence on the solubility of atPAD4 after overexpression in *E.coli* cells. No further experiments were conducted in this thesis to recombinantly express atPAD4.

3.2.2. Sequence analysis of PAD4 and EDS1 variants

Both EDS1 and PAD4 are conserved among most higher plants (Baggs et al. 2020), indicating their crucial role in the plant immune system. As the expression of atPAD4 had failed, I looked for PAD4 orthologues from other species that might provide a suitable alternative. It is unknown why atPAD4 is not expressed soluble in *E.coli* while the similar proteins atEDS1 and atSAG101 can easily be expressed in high amounts.

Due to improved DNA-sequencing methods and bioinformatics information about genomic DNA is cheap and easy to achieve. For example, today the sequencing of a whole human genome cost less than a thousand euros and is offered by several commercial companies[12]. Therefore, a plethora of genomes of different species are available in several databases. This allows the prediction of numerous potential EDS1- and PAD4-homologues. To determine if there is a major difference between the sequence of atPAD4 and atEDS1 and its orthologues in other species, sequences from a broad range of different plant groups and families were chosen and compared. The sequences were taken from the NCBI protein-databank or were provided by the Walter Gassmann group (University of Missouri, *V. vinifera*) and Jane Parker group (MPI for plant breeding research, Cologne, *Oryza sativa and Hordeum vulgare*). The alignment was done using the ClustalOmega tool (Sievers et al. 2011) and the BLAST suite. The sequences were selected to cover a broad range of different plants from different groups and families. Figure 15 shows the comparison of twenty-five different EDS1 homologues including *Picea sitchensis* as a member of the pinophyta-division. Several groups of angiosperms are represented including the amborellales (*Amborella trichopoda*), nymphaeales (*Nelumbo nucifera*), magnoliidae (*Cinnamomum micranthum,* PAD4-only) as well as several monocots and eudicots each.

The sequence comparison of the EDS1 orthologues shows that all variants possess an intact catalytic triad except for *A.trichopoda* that is lacking the catalytic histidine. Most sequences show either a serine or a threonine in the sequence directly after the catalytic serine, similar to *A.thaliana* where it has been shown to block the oxyanion hole (Figure 27), preventing the canonical catalytic mechanism for lipases to occur .The exceptions, beside *A.trichopoda* that is lacking the necessary histidine as described above, are *P.sitchensis, C. sativa* and *V.viniferia* where a isoleucine, alanine or tryptophane is found respectively. While the bulky tryptophane might lead to a steric clash, the isoleucine and alanine might allow the presence of a proper oxyanion hole in the respective sequences. Nevertheless, all EDS1 variants show an insertion after the catalytic aspartate similar to atEDS1, where this insertion forms the αF, αG and αH helices that blocks access to the active centre of the protein. In conclusion, all these variants are unlikely to be active lipase like enzymes just like the atEDS1 tested by Wagner (2013). The sequence identity compared to atEDS1 ranges from 31.08% for *P.sitchensis* to 86.36% for *Arabidopsis lyrata*, all proteins consist of a N-terminal lipase-like domain and a C-terminal EP domain. No obvious difference between the analysed sequences was found, that would point to a distinct difference between the individual EDS1 homologues.

12 https://nebula.org/whole-genome-sequencing-dna-test/

Figure 15: Sequence comparison of EDS1 homologues from several plants. Sequences were taken from the NCBI databases or from sequencing results of cooperation partners. The sequences were aligned using the ClustalOmega server. The residues of the catalytic triad are marked with a black box. The insertion not found in canonical α,β hydrolases is marked with a red box.

Figure 15 shows the alignment of the PAD4- like sequences including the same species as the EDS1-alignment excluding *A. trichopoda* that does not possess an PAD4 homologue (Baggs et al. 2020) but including *C. micranthum* from the magnoliidae group. The alignment of the sequences shows that all twenty-five proteins possess an intact catalytic triad and therefore the potential to possess a hydrolase activity. Unlike the EDS1-sequences most PAD4 homologues do not have a polar amino acid immediately following the catalytic serine. The exception being the member of the Brassicaceae family, including *A.thaliana*, where a threonine follows the catalytic serine and therefore potentially blocks the oxyanion hole. Another difference between the *Brassicaceae* and the other species is the lack of an insertion after the aspartate of the catalytic triad. While the other PAD4 homologues do possess an insertion similar to the one found in atEDS1, all compared *Brassicaceae* sequences lack this insertion, that is blocking access to the active site in atEDS1. The sequence identity compared to atPAD4 ranges from 29.61% for *P.sitchensis* to 87.64 % for *A.lyrata*.

Figure 16: Sequence comparison of PAD4 homologues from several plants. Sequences were taken from the NCBI databases or from sequencing results of cooperation partners. The sequences were aligned using the ClustalOmega server. The residues of the catalytic triad are marked with a black box. The insertion not found in canonical α,β hydrolases is marked with a red box. Species selected that belong to the Brassicaceae family of plants are labelled in a green box. These plants lack the insertion in the N-terminal domain found in members of other families as well as in all EDS1 homologues analysed.

As, regarding the PAD4-homologues, a clear difference between sequences taken from Brassicaceae to other species exists, several variants were tested for soluble expression in *E.coli.*

3.2.3. Expression of different PAD4 homologues

The expression of his-tagged PAD4 variants was tested in several different *E.coli* strains. PAD4 homologues from *Arabis alpina* (aaPAD4), a member of the Brassicaceae family and *Solanum tuberosum* (stPAD4, Potato) were already tested by Stephan Wagner (2013) but are included, together with atPAD4, for completeness and redundancy's sake. Additionally, PAD4 from *Hordeum vulgare* (barley, hvPAD4), *Oryza sativa* (rice,osPAD4) and *Vitis viniferia* (wine grape, vvPAD4) were tested. The result of these expressions is shown in figure 17. For stPAD4, hvPAD4 and osPAD4, a clear overexpression of a protein the right size is detectable after induction with IPTG, unfortunately all three proteins were solely found in the insoluble fraction after centrifugation. In the case of aaPAD4, a slight portion of the protein seems to be in the soluble fraction after lysis and centrifugation. However, this did not lead to a significant amount of soluble protein so that this avenue was not followed further.

Both atPAD as well as the atPAD4-atEDS1 complex expressed using a bicistronic mRNA was found solely in the insoluble fraction. The expression of aaPAD4, osPAD4 and hvPAD4 was drastically enhanced using the Rosetta strains, providing a higher amount of rare tRNAs. The expression of vvPAD4 led to a significant accumulation of protein the right size that also was found to be at least partially soluble (figure 18B). As vvPAD4 proved to be a viable alternative to atPAD4 no further variants were tested in this thesis. Additionally, to the strains shown in figure 17, *E.coli* expressing additional chaperones (Takara 1-5, table 7) were tested for each variant but also proved to be ineffective. Since the expression of osPAD4 and hvPAD4 relies on the presence of tRNAs provided by the pRARE plasmid present in the Rosetta strains, it would be worthwhile to generate a synthetic, codon optimized construct to test against strains with additional chaperones. Additionally, co-expression with the species-specific EDS1 homologue might be successful.

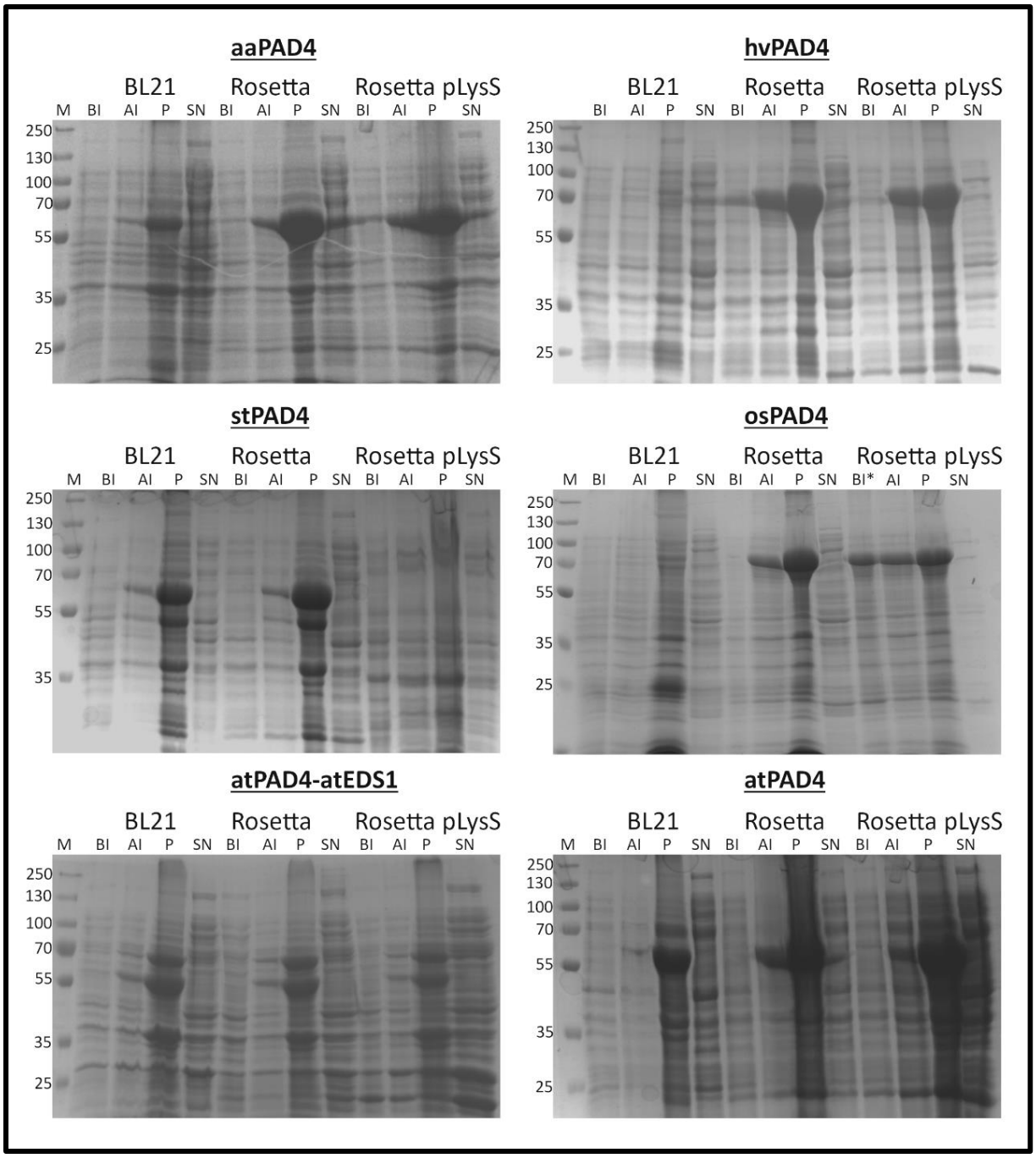

Figure 17: Testexpressions of PAD4 variants, Arabis alpina PAD4 (aaPAD4, 60 kDa), Hordeum vulgare (barley, hvPAD4, 71 kDa), Solanum tuberosum PAD4 (potato, stPAD4, 65 kDa), Oryza sativa PAD4 (rice, osPAD4, 74 kDa), Arabidopsis thaliana (atPAD4, 61 kDa) as well as both Arabidopsis PAD4 and EDS1 in a biscistronic pET-Duet vector (60 kDa and 72 kDa). Beside BL21 (DE3) cells, Rosetta cells with an additional plasmid for rare tRNAs as well as cells expressing the T7 lysosyme (pLysS) were used. Protein expression was induced by addition of IPTG. After lysis of the cells soluble and insoluble parts were separated by centrifugation. M=marker, BI= before induction with IPTG, AI= after induction with IPTG and incubation at 18°C overnight. P= Pellet, insoluble fraction after cell lysis and centrifugation. SN=supernatant, soluble fraction after cell lysis and centrifugation. The before induction sample off Rosetta pLysS for osPAD4 was taken several hours after induction with IPTG and is marked with an asterisk.

3.2.4. Purification of vvPAD4

As the test expressions proved to be successful for vvPAD4, a large-scale expression and purification using the same buffers and protocols as for atEDS1 was attempted. The resulting chromatogram of the gelfiltration is shown in figure 18A. Three distinct peaks could be detected that were further analysed via SDS-gel electrophoresis (figure 18B). The second peak at roughly 60 ml contained a protein of the expected site for vvPAD4 of roughly 70 kDa. The size was further validated using a calibrated analytical gelfiltration column (figure 18C). The third peak at roughly 70 ml contained a smaller protein of roughly 30 kDa that was discarded as an impurity. Interestingly no protein was visible after Coomassie staining for the first peak at roughly 40 ml.

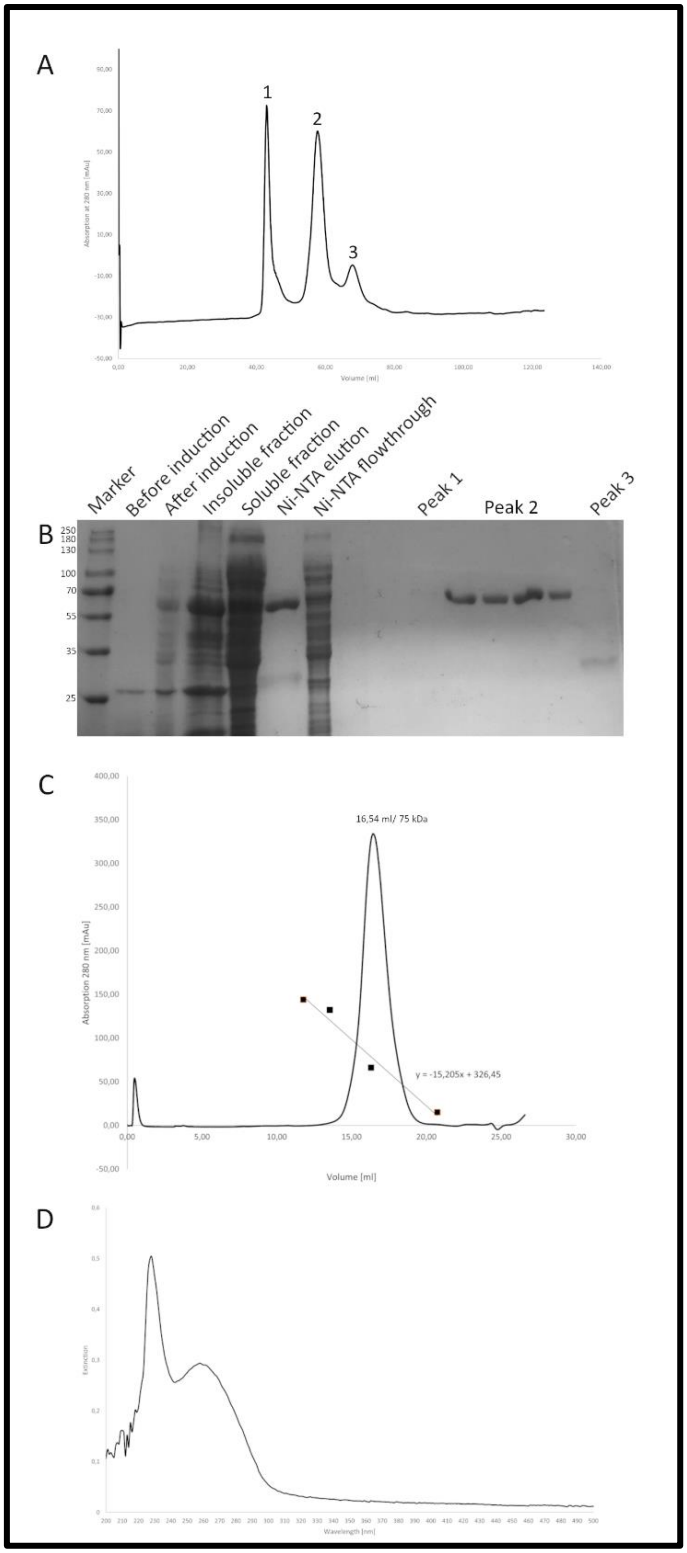

Figure 18: Purification of His-tagged vvPAD4 from E.coli. Figure A shows the chromatogram of the gelfiltration after IMAC using a Superdex75 16/60 column. Three peaks were detected and are labelled. Figure B shows the Coomassie stained SDS-PAGE of the purification. Figure C shows the second gelfiltration of peak 2 on a Superdex200 13/30 column calibrated with proteins of known mass. The calculated mass for vvPAD4 is 75 kDa, indicating a monomeric quaternary structure (vvPAD4 molecular weight= 68 kDa). Figure D shows the UV/Vis absorption spectrum of the pooled Peak 1 as described in port A. The spectrum fits a nucleic acid.

To analyse the content of this peak, that shows a significant absorption at 280 nm, a UV-Vis spectrum was recorded (figure 18D). Two absorption maxima at 230 nm and 260 nm were found. This spectrum fits to a nucleic acid, but not to a protein, which easily explains the lack of band in the Coomassie staining. As the lysis buffer contained DNaseI the nucleic acid is probably an RNA.

The identity of the protein was confirmed via western-blot (using an anti His-tag antibody) and later using mass-spectrometry. The purified protein was then applied to crystallization screens in several concentrations ranging from 1.5-5.0 mg/ml. A higher concentration could not be achieved as the protein started to precipitate. Unfortunately, the protein immediately precipitated in roughly 90 % of the conditions and in no case a crystal formation was observed.

3.2.5. Activity assay

As shown in figure 16 vvPAD4 does have an intact catalytic triad as part of the N-terminal α/β hydrolase domain. To test if vvPAD4 shows lipase-like hydrolase activity, an assay using artificial chromogenic substrates was performed. As a negative control two inactive mutants were created. Both mutants Ser134Ala and His305Ala should show no activity, if indeed a canonical lipase-like activity is catalysed by vvPAD4. The results of the activity assays using para-nitrophenol esters are shown in figure 19. Beside a buffer-only control, lysozyme was used as a further negative control. A lipase from the fungus *Mucor miehei* was used as a positive control. It should be noted that this lipase has a substrate specificity for longer fatty-acid esters and therefore shows very low activity with the used esters.

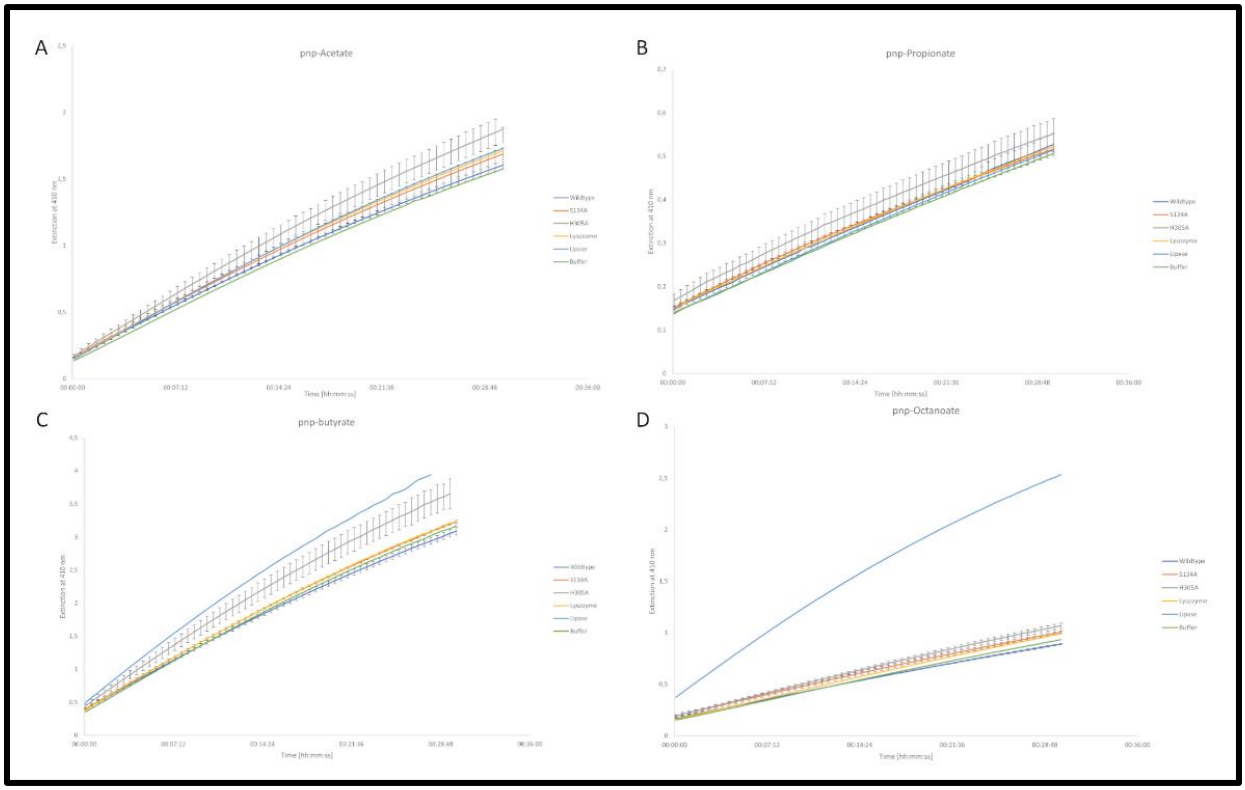

Figure 19: Lipase activity test for vvPAD4 using para-nitrophenol esters of different chain lengths. Wildtype vvPAD4 as well as the catalytic triad mutants S134A and H305A were tested for lipase-like activity. Lysozyme as well as a buffer only sample were used as negative controls. A lipase from Mucor miehei was used as a positive control. The lipase has a specificity for longer chained substrates and therefore shows low activity for the short-chained esters. The release of para-nitrophenolate was measured over time by detecting the extinction at 410 nm. No activity for vvPAD4 was detected. Residual activity of the mutant samples can be explained by impurities in the protein sample.

Using para-nitrophenol-acetate as substrate it was observed that buffer and wildtype protein show the same rate of para-nitrophenolate release. The curve for the lipase, the lysozyme and the Ser134Ala mutant is marginally steeper. The highest release rate was detected for the His305Ala mutant. The positive control shows significant activity only for the para-nitrophenol-octanoate test. Buffer control, lysozyme, the wildtype protein and the Ser134Ala mutant show the same rate of hydrolysis. The His305Ala mutant shows an activity lower than the lipase but higher than the other rows. The same is true when para-nitrophenol-butyrate and para-nitrophenol-octanoate were used, with a higher activity detected for the lipase in each case. While these results point towards a lack of lipase-like activity for vvPAD4, the high background activity due to auto-hydrolysis of the substrate would prevent the detection of low activities. Further experiments especially including different protein and substrate concentration have to be performed.

3.2.6. Buffer optimization for PAD4 purification

To optimize the vvPAD4 purification and achieve a more stable protein solution a variety of buffer substances at different pH were tested in addition to different concentration of NaCl using a thermal shift assay. The results of the screening are displayed in figure 20 A. Generally, the protein showed a higher melting point at 300 mM and 500 mM NaCl in contrast to lower salt concentrations, with the melting point increasing from roughly 38°C to 41 °C. A preference for a single buffer could not be found as a broad range of buffer substances and pH values between 7.5 and 9 resulted in equal results. Therefore, as former purifications were successful using this buffer system, HEPES buffer at pH 8 was further used.

In a second step the influence of several additives was tested to find possible ligands and generally stabilizing agents. The results shown in figure 20B show an improvement of the melting point from 39° to 41.5 °C after the addition of 100 mM NaI. Another thermal shift assay was performed to verify and further analyse the influence of sodium-iodide on the melting point of vvPAD4.

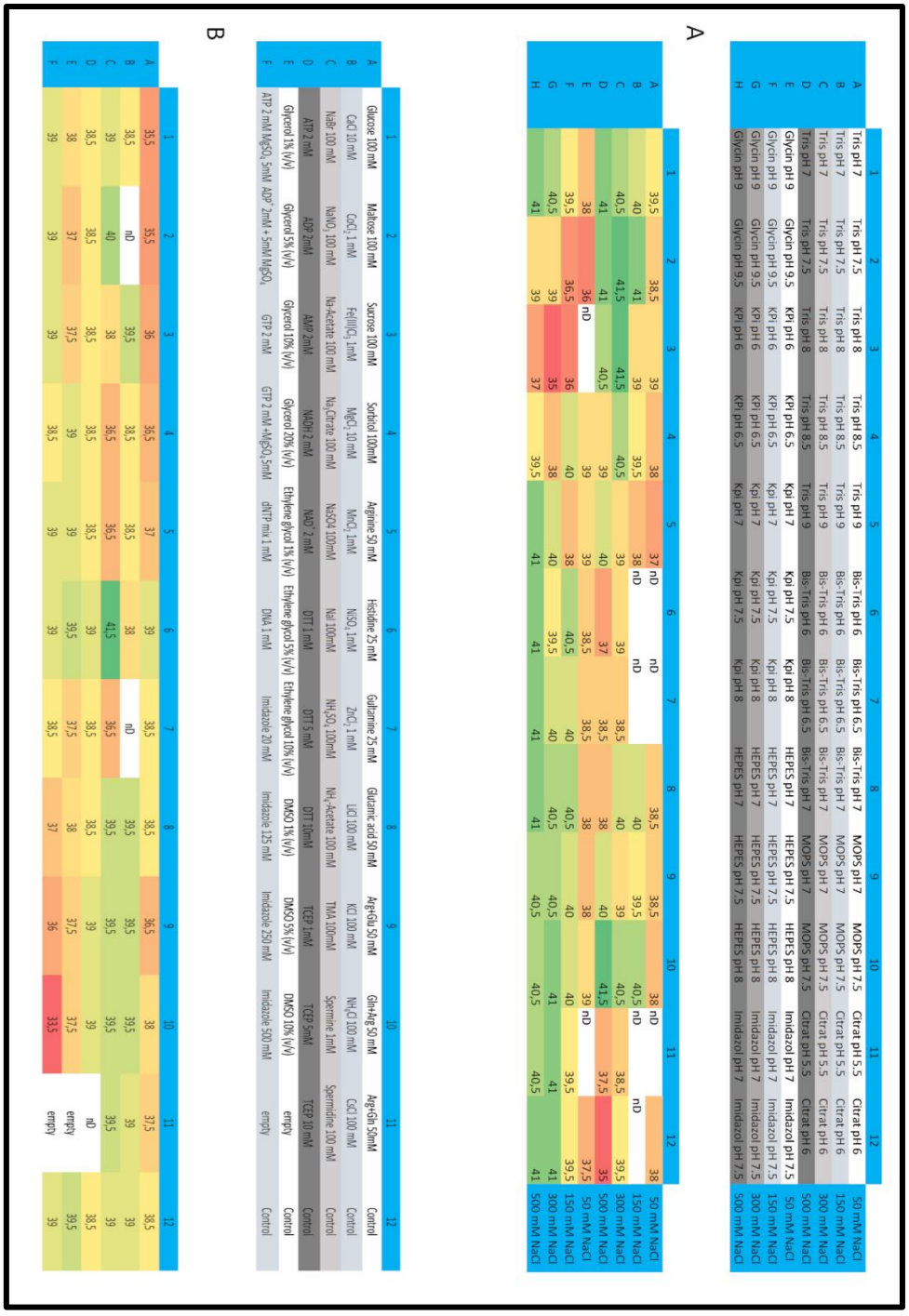

Figure 20: Results of the buffer optimization for vvPAD4 expressed in E.coli. Thermal shift assays using SYPRO orange as dye were performed to optimize the stability of the protein in solution. Shown is the detected melting point colour coded in a heatmap with green for high and red for low melting points. Figure A shows the general screening for pH values, buffer substances and NaCl concentrations. Figure B shows the screening for potential additives uncovering a positive effect of NaI. The melting points were measured in °C.

The stabilizing effect of sodium iodide could be verified. The melting point of vvPAD4 increased gradually with increasing sodium iodide concentrations. For example, a sample in buffer containing 150 mM NaCl showed a melting point of 42°C. Addition of sodium iodide led to a final melting point of 45.5 °C at 300 mM additional NaI. The increase in melting point cannot be explained just by the increase of ionic strength. For example, was the melting point with 300 mM NaCl detected to be 41 °C while a sample with 300 mM sodium iodide had a melting point of 45.5 °C. This indicates a stabilizing impact of iodine ions on vvPAD4, an effect that might even be useful for the collection of anomalous x-ray diffraction data after crystallization. These results are displayed in figure 21.

	1	2	3	4	5	6	7	8	9	10
A	no salt	50 mM I	100 mM NaI	150 mM NaI	200 mM NaI	250 mM NaI	300 mM NaI	500 mM NaI	1 M NaI	300 mM NaCl no I
B	0 mM I	25 mM I	50 mM I	100 mM I	150 mM I	200 mM I	250 mM I	300 mM I	500 mM I	300 mM NaCl no I
C	0 mM I	25 mM I	50 mM I	100 mM I	150 mM I	200 mM I	250 mM I	300 mM I	500 mM I	300 mM NaCl no I
D	no salt	50 mM I	100 mM NaI	150 mM NaI	200 mM NaI	250 mM NaI	300 mM NaI	500 mM NaI	1 M NaI	300 mM NaCl no I
E	0 mM I	25 mM I	50 mM I	100 mM I	150 mM I	200 mM I	250 mM I	300 mM I	500 mM I	300 mM NaCl no I
F	0 mM I	25 mM I	50 mM I	100 mM I	150 mM I	200 mM I	250 mM I	300 mM I	500 mM I	300 mM NaCl no I

	1	2	3	4	5	6	7	8	9	10
A	39	42	43,5	44	44,5	45	45,5	45,5	45,5	42,5
B	43	43,5	44	44,5	44,5	45	45	45	45,5	42,5
C	nD	42,5	43	44	44	44,5	44,5	45	45,5	42,5
D	38,5	42	43	44	44,5	45,5	45,5	46	45,5	41
E	42	42,5	43,5	44	44	44,5	44,5	45,5	45	40,5
F	45	42,5	43	44,5	44,5	45	45,5	45,5	45,5	41

Figure 21: Results of thermal shift assays of vvPAD4. The influence of several concentrations of NaI and NaCl were tested. Shows is the melting point in °C in a heatmap using green for high and red for low melting points. All salts were solved in a buffer containing 50 mM HEPES pH=8. Lanes A-C and D-F contain protein from a different purification. The protein used for lane D-F was purified two days before the batched used for lane A-C. Sodium iodide shows to have a stabilizing effect on vvPAD4.

In accordance to these experiments the buffers were adjusted to include 300 mM sodium iodide and the purification was repeated. As a result, the protein could be concentrated to higher concentrations without resulting in precipitation. The crystallization screening was repeated using these new buffer conditions with concentrations up to 15 mg/ml. Unfortunately, no crystals were found even after several weeks of incubation at 4 °C.

3.3. Expression and crystallization of vvEDS1

3.31. Expression of vvEDS1

As the crystallization of vvPAD4 on its own failed, the next step was to generate and purify a vvPAD4-vvEDS1 complex. To achieve this, the gene for vvEDS1 provided by Walter Gassmann (University of Missouri) and noted as verified, was cloned into the pETM11 vector for expression in *E.coli*. The expectation was to express and purify the roughly 70 kDa protein vvEDS1 equipped with an N-terminal His-tag.

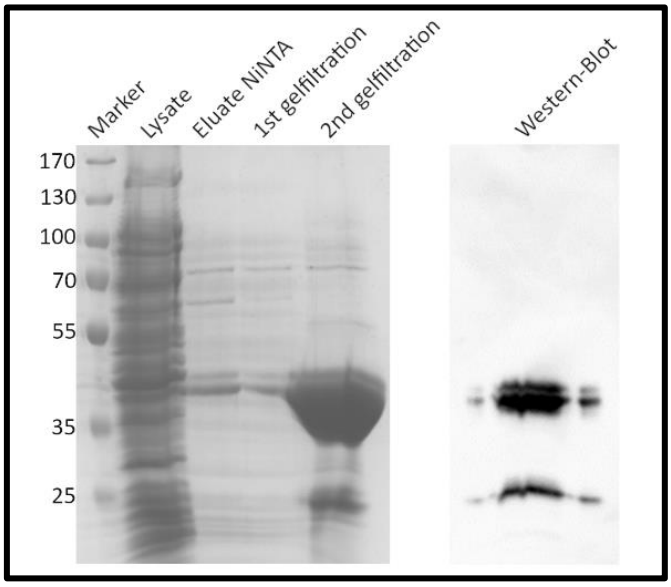

Figure 22: Coomassie stained SDS PAGE and western blot of the purification of vvEDS1^{Nterm}. The resulting protein after purification is shown to be roughly 40 kDa of size instead of the expected 70 kDa for a full lengths vvEDS1. The western blot using an anti-Histag antibody proves the purified protein does possess a His-tag, proving that the purified protein is not an impurity.

As shown in figure 22 unexpectedly a roughly 40 kDa large protein was the result after the purification. Western blotting proved that the purified protein indeed had a His-tag and therefore was not an impurity.

3.3.2. Sequencing of the vvEDS1 construct

To distinguish whether the shortened product is a result of a protease cleavage, a mutation or error in the cloning process, the plasmid was sequenced. To cover the whole sequence, primers sitting in the middle of the sequence were designed.

Figure 23: Sequence comparison of vvPAD4 sequences. Figures A and B show the genomic sequences, taken from the NCBI database, the sequencing result of the sequence as cloned in the expression vector for E.coli expression, as well as the expected sequence as communicated by our cooperation partner. The expected and sequenced sequences were taken from a Cabernet Sauvignon Vitis vinifera plant, while the NCBI genomic sequence stems from a Pinot Noir plant. Several silent point mutations were manually corrected to match the Cabernet Sauvignon sequence. The gene contains two introns in its sequence, these parts are shown in figures A and B. While the first intron was correctly removed in the splicing process, the second intron was retained in the DNA provided by the Gassmann group (University of Missouri). Figure C shows the resulting protein sequences, first the expected sequence resulting from a completely spliced mRNA and the second sequence as a result of the intron retention shown in figure B.

The results show that an unexpected insertion that includes a stop codon and results in a 42 kDa protein is the reason for the shortened product (figure 23). A BLAST search of the insertion, proved that the insert is part of the genomic DNA, but not of the expected mRNA, identifying the sequence as an intron. Further analysis of the whole gene showed that another intron upstream was correctly removed in the splicing process (figure 23B). This indicates that the shortened protein is the result of an alternative splicing event, namely an intron retention. As the original sequence was taken from a cDNA library, this indicates that the resulting protein is indeed biological relevant. In the moment no further information of the ratio and occurrence of this particular alternative splice are available.

3.3.3. Interaction of vvEDS1^{Nterm} and vvPAD4

In a next step vvEDS1^{Nterm} (42 kDa) and vvPAD4(70 kDa) were mixed in equimolar amounts, incubated for thirty minutes on ice and applied to a calibrated gelfiltration column. Additionally, vvEDS1^{Nterm} was applied separately to determine the oligomeric state. The results displayed in figure 24 show one single peak for both experiments. A calibration of the column with known proteins let to a calculated mass for the vvEDS1^{Nterm} of 77.9 kDa and for the mixture with vvPAD4 for 71.5 kDa. This indicated that vvEDS1^{Nterm} does not bind vvPAD4, but forms homodimers.

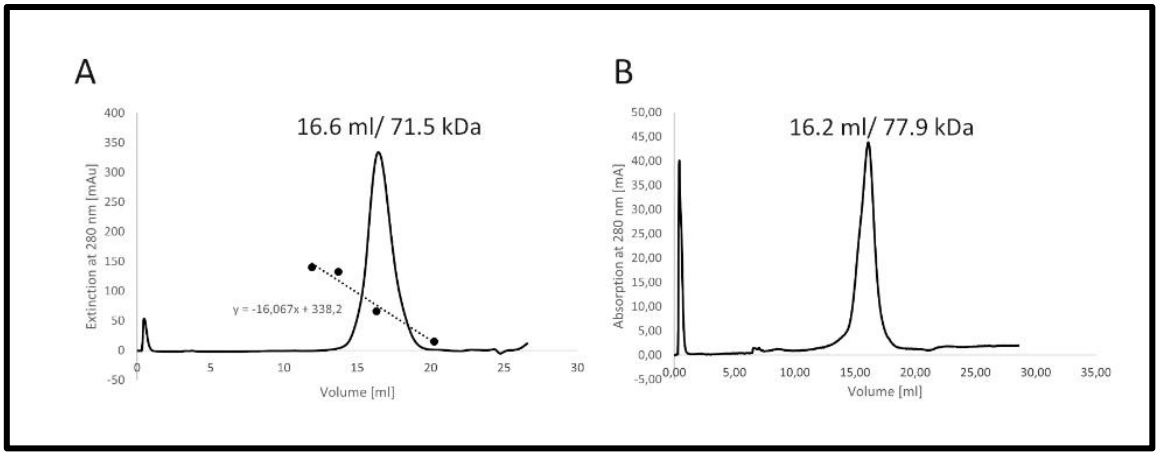

*Figure 24: Gelfiltration experiments of vvEDS1^{Nterm} alone and mixed with vvPAD4. Figure A shows the result after thirty minutes incubation of equal molar amounts of vvPAD4 and vvEDS1^{Nterm}. A single peak consistent with a vvPAD4 monomer (70 kDa) and a vvEDS1^{Nterm} homodimer (42 kDa*2) was found. Figure B shows the result of vvEDS1^{Nterm} alone on the same column. Again, a single peak fitting a homodimer was found. The N-terminal part of vvEDS1 does not bind to vvPAD4.*

3.3.4. Crystallization of N-terminal domain of vvEDS1

While no vvEDS1-vvPAD4 complex could be purified, the N-terminal vvEDS1 was applied to crystallization screens at a concentration of 4.5 mg/ml. Crystal formation was observed after 3-10 days incubation at 4 °C. Most crystals appeared in a shower of often intergrown plates with some conditions providing larger crystals. The protein crystallized in a broad range of conditions containing mostly 15-25% PEG3350, a pH value between 6.5 and 8.5, and about 0.2 M of different salts including, sodium-sulphate, citrate, nitrate, iodide, bromide, fluoride, tartrate, and malate. Optimization of the conditions and upscaling of the drop sizes led to crystal growth after 1-10 days incubation at 4 °C. Some representative pictures of the crystals are shown in figure 25. Several crystals were mounted, frozen and shipped to a synchrotron for diffraction measurement. The best diffracting crystal was grown in a drop of 3 µl containing protein and reservoir solution in a 1:1 ratio. The reservoir consisted of 0.3 M CsCl, 100 mM Bis-Tris pH 8.5 and 12 % PEG3350.

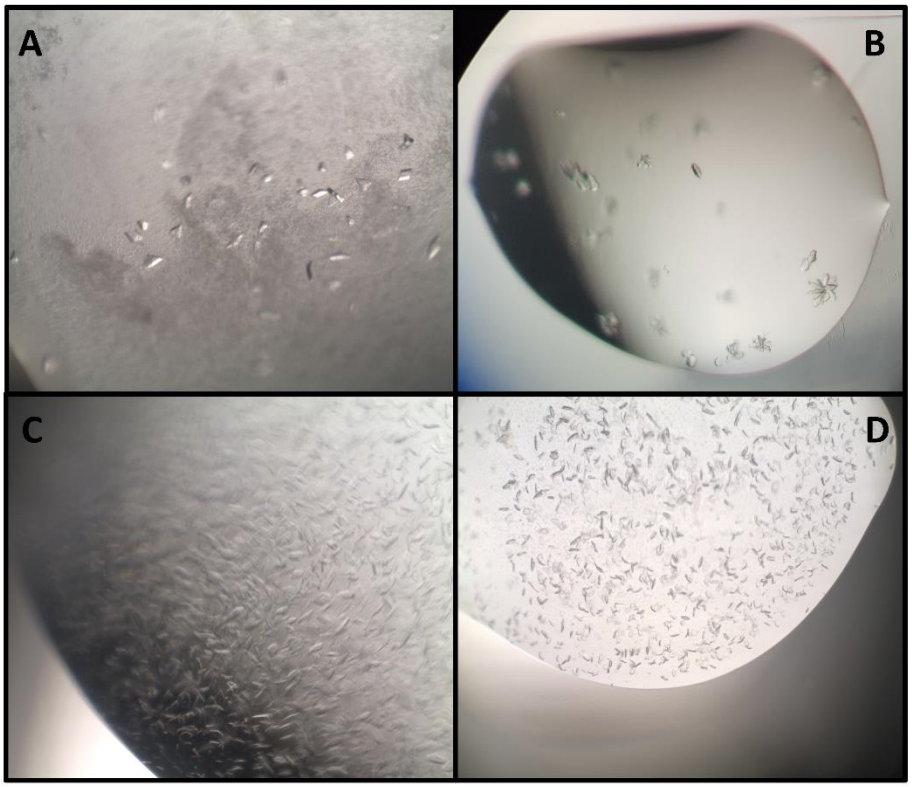

Figure 25: Representative crystals of the vvEDS1^{Nterm} protein, found after 1-5 days incubation at 4 °C.

3.3.5. Molecular replacement and refinement of the vvEDS1^{Nterm} structure

Diffraction measurement at the ESRF (Grenoble, France), allowed the recording of a dataset for the N-terminal vvEDS1 crystal at a resolution of 1.7 Å. The dataset was processed using the Autoproc pipeline in the spacegroup C2 at a resolution of 43.8-1.74 Å. The phase problem was solved using molecular replacement with the N-terminal domain of atEDS1 taken from the atEDS1-1AT73 structure. A first model was then built in the electron density using the Autobuild program of the Phenix (Liebscher et al. 2019) suite. This was followed by several rounds of manual modelbuilding and refinement. The relevant statistics are shown in table 17. The structure was not yet deposited in the PDB at the time this thesis was written.

Table 17: Statistics and data of the data collection, processing and model building of vvEDS1^{Nterm}

vvEDS1-N-terminus		
X-ray diffraction data collection		
Wavelength [Å]		0,91508
Synchrotron and beamline		ESRF/Grenoble ID23-1
Space group		C2
Unit cell	a, b, c [Å]	144.549, 65.077, 45.454
	α, β, γ [°]	90, 105.47, 90
vvEDS1-N-terminus per asymmetric unit		1
Resolution [Å] (highest resolution shell)		43.806-1.742 (1.937- 1.742)
R_{sym} [%]		11.2(68.2)
CC ½		0.982 (0.636)
Signal to noise ratio (1/σ)		5.6 (1.3)
Number of unique reflections		30040 (1503)
Completeness (spherical) [%]		72.1 (13.3)
Completeness (ellipsoid) [%]		89.4 (45.2)
Multiplicity		2.7 (2.8)
Wilson B-Factor [Å^2]		23.58
Structure refinement and validation		
Number of reflections for R_{work}/R_{free}		30037 (102)
R_{work}/R_{free} [%]		0.17/0.21
Number of non-H-atoms		3004
Protein		2689
Water		267
Ethylene glycol		120
Average B-factor [Å^2]		34.67
Protein		33.9
Water		37.86
Ethylene glycol		48.91
RMS deviations:		
Bond lengths [Å]		0.015
Bond angles [°]		1.29
Ramachandran plot		
Favoured <-.(%)		35.15
Allowed (%)		4.24
Outliers (%)		0.61

4. Discussion

4.1. Optimization strategy for protein expression and stability

4.1.1. Expression and purification of atEDS1

One of the main problems I encountered while working with purified atEDS1 was the lack of long-term stability when in solution. While the expression and purification of atEDS1 from *E.coli* reliable resulted in yields of 2-3 mg/l culture, on the condition that the protein was kept cold at all times, purified protein could not be stored longer than one day without visible precipitation in the vial. Christine Toelzer and Stephan Wagner as former PhD students working at the same project have optimized the buffer system used and established the buffers used in this thesis as optimal. Beside a low salt concentration and a pH over 7.5, the addition of fresh DTT or other reductive agents is essential for the stability of atEDS1 in solution. As shown in figure 13B, the continued addition of fresh DTT can prevent aggregation of atEDS1 completely, even over a course of four days. Additionally, the lack of long-term stability in solution might be a reason why most of the attempts to upscale protein crystallization experiments failed. Larger drops require more time to equilibrate with the reservoir solution, in these cases aggregation might occur faster than crystallization.

Although no explicit experiments were performed to verify that disulphide bridge formation is the driving force of aggregation for atEDS1, the experiment shown in figure 13B strongly suggests that indeed a reductive environment is essential for atEDS1 stability in solution. In general, disulphide bridges are formed between two cysteines and occur in an oxidative environment. In total, the atEDS1 sequence contains eight cysteines across the protein. An analysis of their positions in the atEDS1-1AT73 structure shows that Cys73, Cys153, Cys298, Cys399, Cys544 and Cys547 are buried in the protein without contact to the surface (figure 26A). These residues are therefore unlikely to be the cause for the aggregation of atEDS1 in an oxidizing environment. Cys442 is located at the end of the αP helix and is somewhat buried by the αH helix. As the molten state of the helix, found in the atEDS1-1AT73 structure(figure 36), indicates a high degree of flexibility in this region, it is possible for Cys442 to be exposed to the solvent and therefore prone to form disulphide bridges, although a steric clash seems likely (figure 26 B+C). The Cys245 is part of the αG-helix and is directly found at the surface of the protein with the sidechain exposed to the solvent (figure 26B+C). Thus, Cys245 is the prime candidate to be the problematic residue that causes aggregation.

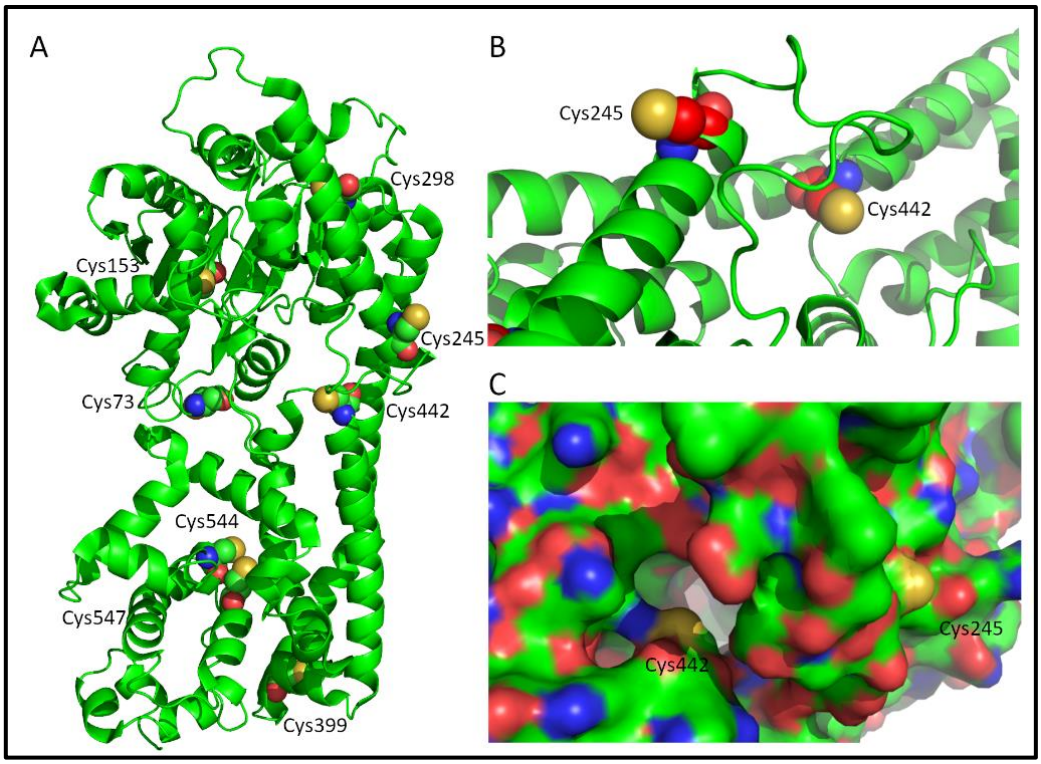

Figure 26: Position of the cysteine residues of atEDS1 taken from the atEDS1-1AT73 structure. Figures A and B show the structure of atEDS1 in the cartoon representation with the cysteines shown as spheres. Figure C shows the surface representation of the area around Cys442 and Cys245 with different colour for different elements. Carbon is represented in green, oxygen in red, nitrogen in blue and sulphur in yellow. Only cysteines 442 and 245 are exposed to the surface of the protein, all other cysteines are buried in the protein core and are not accessible by the solvent.

Due to time restrictions, and as the question of the influence of cysteines on the protein stability was not the central question of this thesis, no further experiments were performed. A quick and easy method to verify the involvement of disulphide bridges in the aggregate formation of atEDS1 would be to do an SDS-PAGE using sample buffer without DTT or other reductive agents. This should, if disulphide bridges between atEDS1 molecules are formed, show bands of higher molecular weight for atEDS1 compared to samples treated with reductive agents. Alternatively, the exposed cysteines could be derivatized using, for example, iodoacetamide an agent that alkylates cysteine residues. The aggregation could be monitored using, as an alternative to gelfiltration, dynamic light scattering (DLS) to determine particle sizes. To avoid the necessity of a derivatization step, critical cysteine should then be mutated to serine an amino acid of similar size to cysteine.

In the last years, as described in the introduction (see 1.2.3.), several atEDS1 interaction partners have been described in literature. While a direct interaction was usually confirmed using Y2H methods, no data about binding constants or kinetics are available. In fact, no published data regarding *in vitro* experiments about the binding of atEDS1 to any proteins is available beyond the published atEDS1-atSAG101 complex (Wagner et al. 2013).

EDS1 and its interactors PAD4 and SAG101 are central and essential mediators between the receptors associated with the immune system and the transcriptional reaction to infection (Dongus & Parker 2021).

Therefore, further studies of interaction between members of the EDS1-protein family and their interactors are of high interest. The generation of a robust, reliable and easy to handle system to characterize and quantify these interactions *in vitro* would be a boon to broaden our knowledge of the intricacies of the plant immune system. Therefore, further experiments with the aim to create a more stable mutant of atEDS1 and subsequently atSAG101 and the atEDS1-atSAG101 complex would be a worthwhile endeavour.

4.1.2. Searching for viable PAD4 orthologues for recombinant expression.

Beside the EDS1-SAG101 complex, the EDS1-PAD4 complex is critical for the function of the plant immune system (Louis & Shah 2014). It can be argued, that investigating the role of PAD4 and its complexes in plant immunity is of a higher social and economic interest compared to the role of SAG101, as monocots do not possess a SAG101-like protein (Lapin et al. 2020). The clade of monocots includes some of the most cultivated plants including rice, corn, wheat, and millet and is therefore of special interest and importance for the world's food supply. Unfortunately, as of yet, very few data are available for PAD4 *in vitro*. In the case of this thesis the main problem is the lack of soluble protein when atPAD4 is expressed *in vitro*, despite several attempts in this thesis (figure 14) and previous (Wagner 2013, Klimpel 2014). To gain soluble protein two main attempts come to mind, either use a different expression system, for example insect cell culture, or try a PAD4 orthologue from another organism for expression in *E.coli.* In this thesis I have chosen to follow the latter strategy.

Due to increasing availability of sequenced genomes and transcriptomes, coupled with the decreasing prices for synthetic genes, the selection of available PAD4 and EDS1 orthologues is vast. The attempt to select sequences based on differences in sequences is shown in figures 16 and 17. I compared twenty-five PAD4 or EDS1 orthologues from a broad evolutionary background of plants. As described in 3.3.2. beside some finer points that might influence a potential catalytic activity of PAD4 and EDS1, the only obvious difference consists of the lack of insertion in the N-terminal domain of *Brassicaceae* regarding the PAD4 orthologues. Thus, no selection can be made based on the analysed sequences alone.

A search of the current literature for pre-existent data regarding PAD4 and EDS1 orthologues revealed some work done *in planta* for both EDS1 and PAD4 from *Vitis vinifera* and *Vitis aestulavis* (Gao et al. 2010, Gao et al. 2014). Additionally, sequences for rice and barley PAD4 as monocots were tested for expression in *E.coli* beside aaPAD4 and stPAD4.

4.1.3. Establishing and optimizing the expression of vvPAD4.

As shown in figure 17 and 18, only the vvPAD4 variant tested in this thesis was found to be expressed soluble in *E.coli*. A typical chromatogram and associated SDS-PAGE are shown in figure 18B and C. After IMAC and gelfiltration vvPAD4 was gained in high purity (figure 18B and C) and satisfying amount for crystallization experiments and biochemical analysis. A second gelfiltration using a calibrated column determined a calculated molecular mass for vvPAD4 of 75 kDa (figure 18 C), that fits to a monomeric quaternary structure (MW vvPAD4= 68 kDa). This finding is consistent with BiFc assays performed in *A.thaliana* (Gao et al. 2014 supplementary information). Interestingly, all gelfiltration experiments after IMAC resulted in three peaks (figure 18A), while one peak consisted of a protein that was easily separated from vvPAD4 in the gelfiltration the third peak contained a nucleic acid as indicated by the corresponding UV/VIS absorption spectrum (Figure 18D). The lysis buffer used in the purification runs was supplied with DNaseI, therefore is likely that the nucleic acid in question is an RNA. While this is an interesting finding, no further effort was made to identify and sequence the RNA, as the co-purification is likely an artifact of the overexpression in *E.coli*. A first step towards identifying the copurified RNA would be to perform a gelelectrophoresis to decipher if the peak contains a single RNA or a mixture of several types.

The successful and reproducible purification of vvPAD4 allows for characterisation of a PAD4 orthologue *in vitro* for the first time. Attempts at crystal screening failed, as the protein immediately precipitated in almost all conditions. Additionally, vvPAD4 could only be concentrated up to 5 mg/ml before precipitation occurred when using the buffers as optimized for the purification of atEDS1. After using thermal shift assays to test several buffer substances, pH values and salt concentrations (figure 20A) as well as possible additives (figure 20B) an improved buffer containing 300 mM NaI was found to be ideal (figure 21). As a result of the new buffer the protein was concentrated up to 15 mg/ml without visible precipitation. Nevertheless, no protein crystals were observed after several crystal screening experiments.

As the crystallisation of unbound vvPAD4 was not successful, it was attempted to produce a vvEDS1-vvPAD4 complex. To test a chimeric complex consisting of atEDS1 and vvPAD4 was not feasible as BiFc studies in *A.thaliana* showed no interaction (Gao et al. 2014). As shown in figures 22 and 23 unexpectedly, the sequence used contained a new, previously unknown splice variant and not the full-length protein. The truncated vvEDS1^{Nterm} did not interact with vvPAD4 (figure 24). The lack of interaction can be explained using the crystal structure that will be discussed later (see 4.3.4.).

The next step in the project should be to express a full length vvEDS1 either in itself or in complex with vvPAD4. Nevertheless, the expression and purification of a PAD4 orthologue in amounts sufficient for crystallisation experiments is a massive step towards gaining a reliable system to investigate PAD4 and its interaction *in vitro*. This thesis is the first written report of a recombinantly expressed, purified PAD4 orthologue.

4.2. Enzymatic activity of EDS1 and PAD4

4.2.1. atEDS1: enzyme of pseudoenzyme?

The N-terminal lipase-like domain of atEDS1 with its α/β-hydrolase fold has an intact catalytic triad consisting of Ser123, Asp187 and His317. Nevertheless, neither was catalytic activity found with wildtype atEDS1 *in vitro*, nor was a phenotype identified after mutating the whole catalytic triad of atEDS1 (Wagner et al. 2013). Interpreting the structure of the atEDS1-atSAG101 complex, the inactivity of atEDS1 can be explained by three factors. On the one hand, the access to the active site is blocked by the helix αF, on the other hand, the backbone nitrogen of Ser46, that would be part of the oxyanionhole, is blocked by forming a hydrogen-bond with the side chain of Ser124. Additionally, the sidechain of Phe47 occupies the space potentially used to form the covalent acyl intermediate at Ser123 (Wagner et al. 2013). Nevertheless, it is unknown if this situation is influenced by the bound state of atEDS1 in the complex structure or if conformational changes allow for access to the active site in an unbound state of atEDS1. The importance of the potential catalytic triad is stressed by the fact that all three amino acids are widely conserved among the plant kingdom for orthologues of EDS1 (figure 15). This indicates that the presence of these three residues is important for EDS1 to function correctly.

The successful crystallization and structure determination of the atEDS1-nanobody complexes allows for an analysis of the active site while not bound to atSAG101. The resolution of the atEDS1-1AT73 (table 16) structure permits the confident placement of the side chains of amino acids in the core of the protein. Due to the low resolution of the atEDS1-1AT15 (table 14) structure and the atEDS1-1AT21 structure (table 15), no conclusion can be made from these structures regarding such fine details. Analysing the catalytic core, all amino acids were placed the same way as in the atEDS1-atSAG101 structure, including Ser124, Ser46 and Phe47. In addition, the helix αF blocks access of any substrate to the active site identically as observed in the atSAG101 bound state (Wagner et al. 2013). While, of course, a conformational change might occur to allow access for a potential substrate, in neither of my three atEDS1-nanobody structures nor in the atEDS1-atSAG101 structure a change occurs in the overall conformational line-up of the protein, thus suggesting atEDS1 to be rather rigid. These results are in accordance with the current state of the literature that classifies atEDS1 as a kind of pseudoenzyme (Wagner et al 2013, Voss et al. 2019, Lapin et al. 2020, Bhandari et al. 2019).

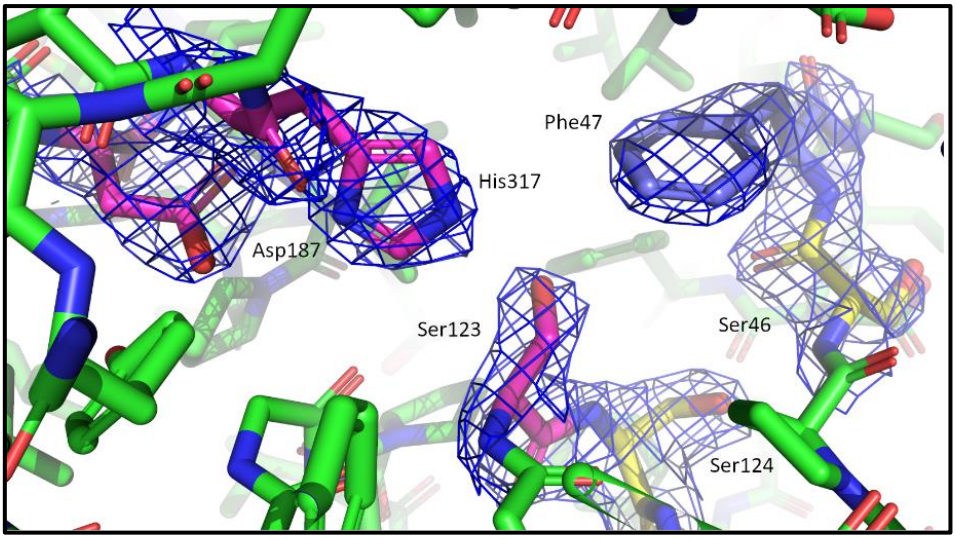

Figure 27: Detailed view of the catalytic triad of the crystal structure of atEDS1 bound to 1AT73. For selected key residues the electron density map is displayed at contour level σ=1. While the catalytic triad is intact, the active site is occupied by Phe47, blocking the space of a potential substrate. Additionally, Ser124 blocks the potential oxyanionhole by forming a hydrogen bridge with the backbone nitrogen of Ser46.

4.2.2. Analysis of the active site of vvEDS1

The structure of the N-terminal domain of vvEDS1 allows the comparison of two orthologues of the same protein regarding the potential active site. Notably, it is reported that vvEDS1 can compensate a loss of atEDS1 after infection with powdery mildew (Gao et al. 2014). Like all analysed EDS1 orthologues, vvEDS1 does contain an intact catalytic triad (figure 15) and has therefore the potential to be an active enzyme. The high resolution of the vvEDS1-N-terminal domain structure (Table 17) allows for the confident placement of all amino acids involved in building the active site. A detailed view of the involved amino acids is shown in figure 28.

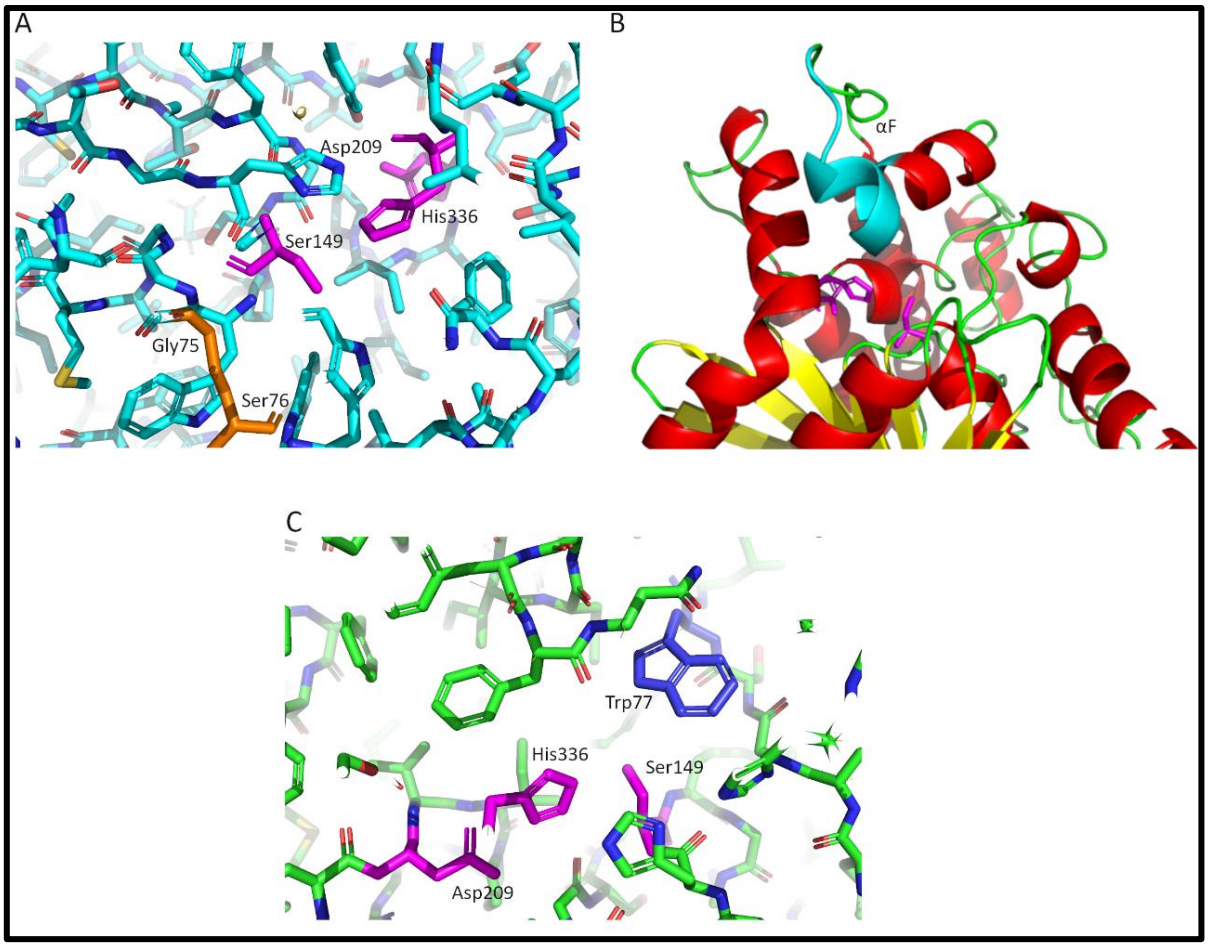

Figure 28: Details of the catalytic triad of vvEDS1. Figure A shows the catalytic triad in magenta as well as a part of the residues forming the oxyanionhole in orange. The catalytic triad is intact and the oxyanionhole unblocked, in contrast to atEDS1 where a hydrogen bridge is preformed to the equivalent of the amino acids represented in orange. Carbon is represented in cyan, nitrogen in blue and oxygen in red, for all other amino acids. Figure B shows the protein in the cartoon representation. The catalytic triad is shown in the sticks representation and is coloured magenta. The helix αF is represented in cyan, the helix blocks access of potential substrates to the active site. Figure C shows the active site with the catalytic triad in magenta as well as Trp77 in blue. The tryptophane occupies the space a potential acyl-intermediate would take up.

Regarding atEDS1, three factors were identified that explains the lack of potential activity (Wagner et al. 2013). Firstly, the access to the active site is blocked by the helix αF, the same is true for vvEDS1 as shown in figure 28B, identical to atEDS1 access to the active site is physically blocked, thus preventing entrance for any potential substrate. The second obstacle explaining the inactivity of atEDS1 was the presence of a phenylalanine residue that takes in the space needed at the catalytic serine to form the acyl-intermediate (figure 27). In the vvEDS1-N-terminal domain structure this space was taken by Trp77 (figure 28C). Therefore, the space needed to form an acyl-intermediate at the catalytic serine is not available for a potential substrate. Both points indicate vvEDS1 being a pseudoenzyme comparable to atEDS1. The third factor explaining the lack of catalytic activity of atEDS1 is the partially blocked oxyanionhole, as Ser124 forms a hydrogen-bond with the backbone nitrogen of Ser46.

This is not the case in the vvEDS1 structure (figure 28A). Figure 15 shows the sequence comparison of twenty-five EDS1 orthologues taken from a broad spectrum of plants.

Most sequences analysed possess either a serine or a threonine directly after the catalytic serine, indicating that these proteins show a similar blocked oxyanionhole. Nevertheless, several sequences show small unipolar amino acids including alanine, isoleucine or leucine following the catalytic serine. The sidechains of these residues are unable to form hydrogen-bonds, and it is therefore likely that the potential oxyanionhole is available in these variants. The vvEDS1 sequence was found to express a tryptophane after the catalytic serine, the only analysed sequence to show such a bulky aromatic residue. This arrangement was found among several *Vitis viniferia* cultivars (Gao et al. 2010, Chong et al. 2008). The sidechain was found to be placed away from the active site, therefore not influencing the potential activity. The structure of vvEDS1 shows that, in contrast to the atEDS1 structure, the oxyanionhole is not implicated by a preformed hydrogen bond, therefore allowing it to participate in a potential canonical lipase like hydrolysation.

In conclusion, while all elements necessary for a potential catalytic activity are in place in case of the vvEDS1 structure, the active site is still inaccessible. Identically to the atEDS1 structure (Wagner et al. 2013) and the atEDS1-nanobody complexes described above, the helix αF blocks access to the active site for all potential substrates while additionally a bulky sidechain blocks off the catalytic serine. Both points make it likely that vvEDS1 is not an active lipase-like enzyme, comparable to atEDS1. Nevertheless, the N-terminal domain of vvEDS1 was easy to crystallise and looked, in a first impression of the atomic B-factors, somewhat more stable compared to atEDS1. As it is known that vvEDS1 can replace atEDS1 (Gao et al. 2014) this construct is suited to further search for potential activity. While it seems likely, as described above, that vvEDS1 is indeed an inactive pseudoenzyme, it cannot be excluded that conformational changes upon contact with a substrate or a binding partner allows for access to the active site. In addition, a more stable protein might prove beneficial if low activities that require long reaction times should be found.

4.2.3. Enzymatic activity of PAD4 and its orthologues

4.2.3.1. Experimental evidence of a potential vvPAD4 activity

Just like its family members EDS1 and SAG101, PAD4 possesses a N-terminal lipase like domain with an α/β-hydrolase fold. During evolution, the catalytic triad was conserved among a broad range of species (Wagner et al. 2013, figure 16), indicating a crucial role of these residues for the function of PAD4. Due to the successful purification of recombinant vvPAD4 from *E.coli*, some experiments regarding the potential catalytic activity were performed. Figure 19 shows the result of these first activity assays using artificial, chromogenic substrates.

During the experiments some problems occurred that caused concern. The first of these problems was the high background activity. The test was performed using a pH value of 8, in accordance with the stability assays (figure 20).

The high pH value of the buffer leads to a deprotonation of the released carboxyl acid, thus rendering the reaction irreversible. This resulted in a very high background activity (figure 19), that could have easily masked weak activities of vvPAD4. The second concern is the necessity of adding some measure of organic solvent due to the low solubility of both product and educt in water. In this thesis 10% DMF was used as final concentration in the assay. While no precipitation was observed, it is unknown if the protein can tolerate this condition or not. The third concern that occurred was the potential unspecific activity of impurities. After a first experiment found a low activity for BSA, a literature search revealed several publications of proteins showing activity in assays using short-chained para-nitrophenol-esters, these proteins included: 3-phosphoglyceraldehyde-dehydrogenase (Park et al. 1961), BSA, metmyoglobin and β-lactoglobuline (Ostdal & Andersen 1995), chymotrypsin and insulin (Hartley & Kilky 1953). Additionally, imidazole can also catalyse the hydrolysis of esters (Kirsch & Jencks 1963). It is therefore plausible to explain low activities, as found for the His305Ala mutant on either impurities or residual imidazole not removed during gelfiltration.

The results measured for the wildtype vvPAD4 and the Ser134Ala mutant point towards vvPAD4 not showing a lipase-like activity (figure 19). Nevertheless, the performed experiments are not sufficient to make a definite conclusion. Additional measurements using different protein- and substrate-concentrations are required. The next step to investigate a potential activity should be to establish an assay capable of identifying even very low activities. A first step might be to use methylumbelliferyl esters that release a fluorogenic product after hydrolysation. Alternatively release of fatty acids from lipids could be directly measured after incubation with vvPAD4 using high performance liquid chromatography (HPLC) or gas chromatography (GC) either on its own or coupled to mass spectroscopy. Another alternative is the photometric detection of the fatty acids by converting the released products to copper soaps (Han & Ge 2000).

No experimental study about a possible lipase-like activity of vvPAD4 is available, therefore there is no way to validate the features discussed above indicating that vvPAD4 shows no such activity. Some information is available regarding atPAD4, that shares a high sequence similarity (figure 15) despite the lack of a large insertion between the canonical helix αG and the strand β8. Although it must be noted that while vvEDS1 can replace atEDS1 in *A.thaliana*, vvPAD4 cannot rescue the loss of atPAD4 (Gao et al. 2013). It is known that an intact catalytic triad is not needed for *A.thaliana* immune response against the pathogens *Hyaloperonospora arabidopsidis* (Wagner et al. 2013; Louis et al. 2012) and *Pseudomonas syringae* (Louis et al. 2012).

Contrary, the immune response against *Myzus persicae* (Green peach aphid) is lowered if either the catalytic serine or aspartic acid is mutated to alanine, although no such phenotype is found if the histidine of the potential catalytic triade is likewise mutated (Louis et al. 2012). For successful defence against the green peach aphid the N-terminal domain of atPAD4 on its own is sufficient to mimic the wildtype phenotype (Dongus et al. 2020).

The histidine of the catalytic triad is essential if one suspects a canonical lipase-like activity as found in other α/β-hydrolase fold containing lipases (Rauwerdink & Kazlaukas 2015). This suggests, that while parts of the catalytic triad of atPAD4 are indeed important for a subsection of the atPAD4 initiated immune response, this effect is not based on a canonical lipase like activity. Potential alternatives are discussed below (4.2.4.) for PAD4 and EDS1 together since these points concern both proteins.

4.2.3.2. Sequence analysis of vvPAD4 and other PAD4 orthologues

No structural information for atPAD4 outside of a model published previously (Wagner et al. 2013) is available, furthermore no experimental information about orthologues of atPAD4 regarding a potential catalytic activity is published so far. Therefore, data concerning the better investigated atEDS1 (Wagner et al. 2013, Voss et al. 2019) need to be taken into consideration. As discussed above, atEDS1 is classified as a pseudoenzyme, because - in spite of an established catalytic machinery - no activity was found *in vitro*, and no phenotype was found *in vivo* after mutating the catalytic triad. One reason identified was the formation of a hydrogen bond between the amino acid following the catalytic serine and a backbone nitrogen of one of the peptide groups making up the potential oxyanion-hole (Wagner et al. 2013). A sequence comparison of PAD4 orthologues (figure 16) shows a distinct difference at this position regarding PAD4 variants from *Brassicaceae* and other plants. The five analysed *Brassicaceae*-sequences all show a threonine following the catalytic serine. This allows the conclusion that the polar amino acid in an arrangement similar to atEDS1 blocks off the oxyanion-hole. This is not the case for any other of the analysed sequences that all show a non-polar amino acid following after the catalytic serine. In this case I expect that the oxyanion-hole is competent to stabilise a potential transition state. This finding is another observation that sets apart *Brassicaceae* PAD4 orthologues from other plants, beside the lack of the critical insertion in the N-terminal domain. Non-*Brassicaceae* sequences analysed all show this insertion that is highlighted in figure 16 in the N-terminal domain comparable to EDS1 (figure 15). What sets the PAD4 and EDS1 orthologues apart is the length of the insertion. The approximately length of the insertions is shown in table 18.

Table 18: Analysis of the length of the insertion in the N-terminal domain of PAD4 and EDS1 orthologues. The results for both proteins were colour coded separately. For C.micranthum and A.trichopoda no EDS1/PAD4 sequence was available.

Species	PAD4 length of insertion	EDS1 length of insertion
Picea sitchensis	47	75
Arabidopsis thaliana	0	77
Arabidopsis lyrata	0	77
Eutrema salsugineum	0	77
Raphanus sativus	0	77
Brassica napus	0	76
Eucalyptus grandis	61	75
Punica granatum	49	81
Cucumis sativus	62	77
Vitis viniferia	59	74
Mucuna pruriens	62	80
Glycine max	62	80
Arachis hypogaea	70	83
Nelumbo nucifera	60	76
Solanum lycopersicum	34	81
Nicotiana tabacum	48	81
Coffea arabica	62	81
Quercus lobata	59	78
Hordeum vulgare	67	79
Oryza sativa	71	79
Zea mays	70	79
Cinnamomum micranthum	66	
Musa acuminata	64	81
Elaeis guineensis	66	82
Phoenix dactylifera	66	83
Amborella trichopoda		84

In the case of EDS1, the insertion length varies between 84 amino acids for *A.trichopoda* to 74 amino acids for *V.viniferia* with atEDS1 (77 amino acids) being in the middle. The structure of the N-terminal domain of vvEDS1 shows a fold almost identical to atEDS1 indicating that 74 amino acids are sufficient to ensure the formation of the αF, αG and αH helix made up by the insertion. In contrast, the insertion found in the PAD4 orthologues is much more variable in length, and ranges from 34 amino acids for *S.lycopersicum* to 71 for *O.sativa*. For vvPAD4, the insertion is 59 amino acids long and thus distinctly shorter than in vvEDS1. Due to the shortened sequence, it is questionable if the insertion is enough to fully form the αF helix that blocks the access to the catalytic triad in vvEDS1 and atEDS1. This might allow a substrate to enter the active site for subsequent hydrolysis.

Unfortunately, all crystallisation experiments performed in this thesis concerning vvPAD4 were not successful, and therefore no experimental structural information is available. Additionally, the experimental data gathered in this thesis are not sufficient to make a definite statement concerning the lipase-like activity of vvPAD4.

To get a first idea about the structures of PAD4 variants, selected PAD4 orthologue protein-sequences were used for structure prediction using the Phyre2 server (Kellay et al. 2015). The model for PAD4 from *Arabidopsis thaliana* was taken from the AlphaFold database (EMBL-EBI, Jumper et al. 2021). While the whole sequences were used for structure prediction, only the N-terminal domains were further analysed. Figure 29 shows the N-terminal domains of the predicted structures of atPAD4, vvPAD4 and the PAD4 orthologue from *Arachis hypogaea* (ahPAD4, peanut) and *Solanum lycopersicum* (slPAD4, tomato). The proteins were chosen for their increasing length of the insertion in the N-terminal domain (Table 18).

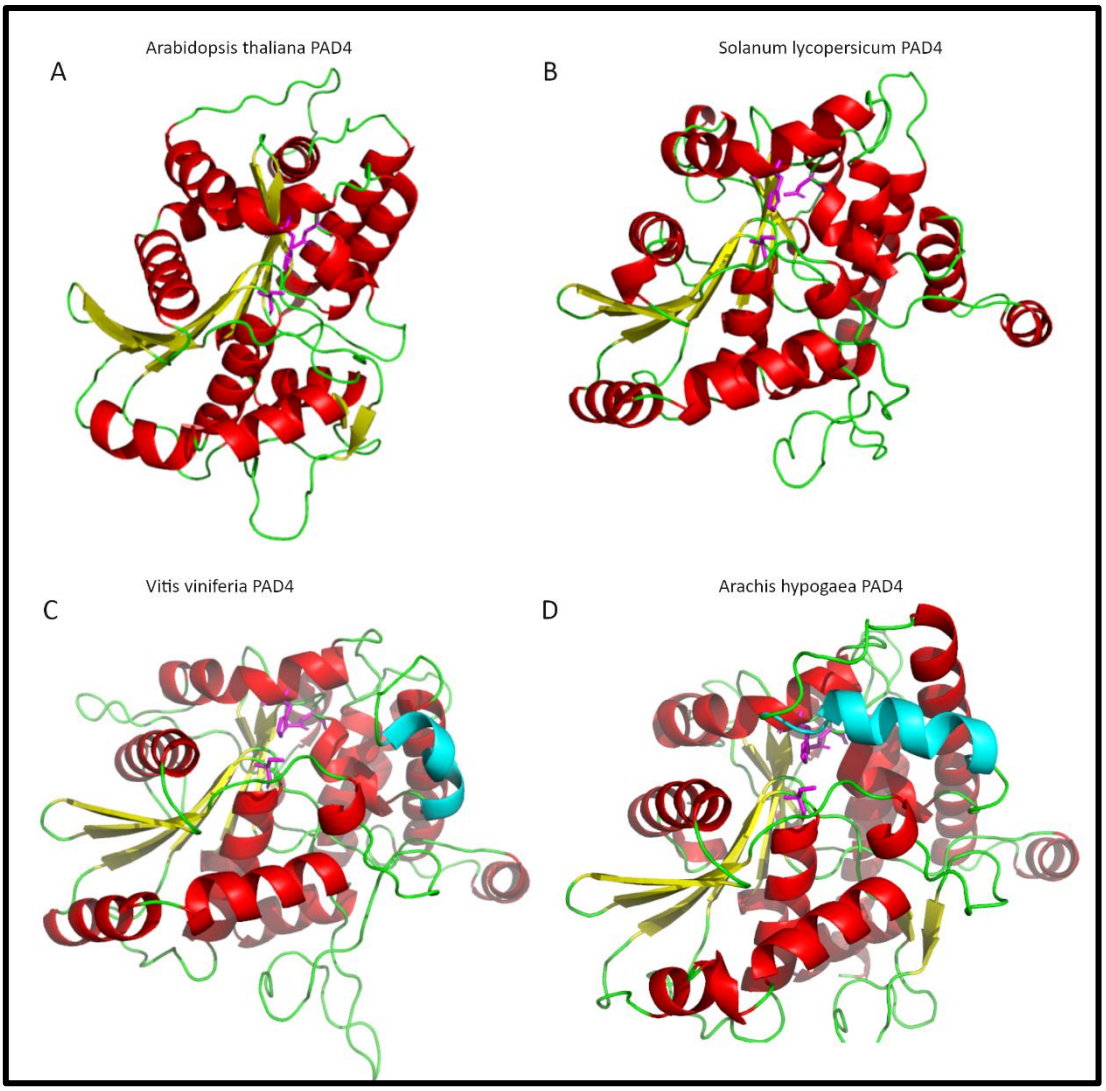

Figure 29: Predicted structures of the N-terminal domain of PAD4 orthologues from different plants. The model for the Arabidopsis thaliana PAD4 was taken from the AlphaFold database while the other models were predicted by the Phyre2 tool. The catalytic triad of each variant is highlighted in magenta while, if appropriate, the helix αF is shown in turquoise. All proteins are represented in the cartoon format with the catalytic triad shown as sticks.

As atPAD4 has no αFGH insertion in its N-terminal domain, the potential catalytic site is accessible to putative substrates. With increasing insertion length more and more of the αF-helix that blocks the active site is formed. Figure 29B shows that, according to the models, slPAD4 active site is accessible comparable to the model of atPAD4 (figure 29A). For vvPAD4 a part of the αF helix is formed but was not placed such that the catalytic triad becomes inaccessible (figure 29C). It has to be noted that the presented structure is indeed a prediction and has to be validated experimentally. The predicted structure for ahPAD4 shows a fully formed αF helix that covers the catalytic triad, comparable to the atEDS1 and vvEDS1 structures.

In conclusion, the question of a potential catalytic activity of PAD4 is hard to answer. The analysis of the sequences shows a clear difference between the *Brassiaceae* and other plants (figure 16).

Published data (Loius et al. 2012, Dongus et al. 2020) proves the importance of a partially intact catalytic triad, but at the same time suggests that not a canonical lipase like mechanism is responsible as the catalytic histidine show no influence. These publications also demonstrate that an intact triad is only necessary for a subset of the atPAD4 function. These data strongly point against a canonical lipase like mechanism. The analysis of sequences and models of PAD4 variants from different species suggests, that for a subset of species the catalytic triad should be accessible for substrates while for others are not, both depending on the size of the insertion in the N-terminal domain.

To answer the arising questions, the next step, besides using the now available vvPAD4 for further experiments, should be to try to overexpress and purify variants with varying insertions e.g., slPAD4, ahPAD4 as well as a member of the *Brassicaceae* family to assess the impact of the insertion on a potential activity. Another avenue to pursue should be to use the known dependency of an intact catalytic triad for *Arabidopsis thaliana* to defend against the green peach aphid to perform *in planta* experiments. A potentially interesting experiment could be to test the impact of the N-terminal domain of *Glycine max* that has been shown to complement a PAD4 knockout plant in their defence against *Pseudomonas syringae* (Wang et al. 2014). This protein variant shows an insertion length of 62 amino acids (table 18) a comparable length to vvPAD4.

4.2.4. Functions of α/β-hydrolases beyond lipase-like activity.

The conservation of the catalytic triad among a broad background of plant species (figure 15+16) indicates the importance of an intact catalytic triad for both EDS1 and PAD4. The N-terminal domain of both proteins was classified as Lipase_3 upon search in the Pfam-Database (Mistry et al. 2021), a class of lipases that usually prefers triglycerides as substrates. These proteins are part of the α/β-hydrolase fold protein family, a fold found in all three kingdoms of life (Mindrebo et al. 2016).

While EDS1 and PAD4 N-terminal domains share the highest degree of similarity with lipases (Wagner et al. 2013, Falk et al. 1999, Jirage et al. 1999), the hydrolysation of triglycerides is not the only reaction catalysed by members of the α/β-hydrolase fold protein family. Beside several other hydrolase reactions, α/β-hydrolase fold containing enzymes can also act as lyases, transferases and isomerases (Rauwerdink & Kazlauskas 2015). Additionally, several proteins classified as non-catalytic family members are known including thyroglobulin and neuroligin, a protein known for its role in autism spectrum disorder (Lenfant et al. 2012). Potentially PAD4 and EDS1 could therefore either catalyse a reaction beside the hydrolysis of esters, or show no activity at all, but nevertheless need an intact catalytic triad to function. Examples for both cases have been described for proteins found in plants.

While both EDS1 and PAD4 contain an intact catalytic triad, as described above, for EDS1 and *Brassicaceae* PAD4 variants the oxyanion hole seems to be blocked. This prevents a canonical lipase-like activity but is common in both decarboxylases and (s)-hydroxynitrile lyases.

In association with the plant immune system these enzymes release toxic compounds to defend against feeding insects (Rauwerdink & Kazlauskas 2015, Schmidt et al. 2008, Auldrige et al. 2008). *A.thaliana* does not contain cyanogenic glycosides, but nevertheless several of these enzymes are present in the plant (Wäspi et al. 1998, Andexer et al. 2007). Remarkably, the degree of similarity between hydroxynitrile lyases and esterases is very high, it was shown that mutation of as little as two amino acids can convert both enzymes to each other (Nedrud et al. 2014, Padhi et al. 2010). Additionally, it is common to find an insertion in catalytic proteins containing an α/β-hydrolase fold often referred as lid domain (Rauwerdink & Kazlauskas 2015). Therefore, in conclusion, it cannot be excluded that both EDS1 and PAD4 do indeed show enzymatic activity that just has not been identified and tested yet. Especially the PAD4 variants from the *Brassicaceae* family show all requirements to possess a lyase activity.

A second possible explanation for the conservation of the catalytic triad, despite the lack of catalytic activity identified until now, is the possibility of the N-terminal domain acting as receptor or binding site for a small molecule. Several examples for this have been described in the literature. The gibberellin receptor gibberellin insensitive 1 (GID1) from both *A.thaliana* and *O.sativa,* which consists of a classical α/β-hydrolase fold, has been shown to be activated upon binding of a gibberellin in the active site (Murase et al. 2008, Shimada et al. 2008). Both proteins do not show catalytic activity as the catalytic triad is disrupted due to a mutation of the canonical histidine to valine. Another example is the strigolactone receptor strigolactone esterase D14 (DWARF14) that does contain an intact catalytic triad. The protein has been shown to hydrolyse strigolactones, although the reaction is extremely slow with a K_{cat} of 0.12 (1/min) (Seto et al. 2019). Additionally further studies suggests that the binding of the small molecule and not the hydrolysation is important for the protein function (Seto et al. 2019). This is cemented by the finding that the inactive mutant D218A is still able to compensate a loss of the wildtype protein in a strigolactone dependent manner (Seto et al. 2019, Marzec & Brewer 2019). A third example is the karrikin receptor karrikin insensitive 2 (KAI2) form *A.thaliana* that recognizes small molecules found in smoke. While the receptor has an intact catalytic triad, no measurable hydrolysation activity was found. The structure of the receptor in complex with its substrate proves that while the triad is involved in the binding of the substrates, no hydrolysation can occur (Guo et al. 2013).

In summary, it can be concluded, that further research is needed both *in vitro* and *in planta* to distinguish the role of the catalytic triads of EDS1 and PAD4. The successful purification of vvPAD4 allows for the first time to perform activity assays *in vitro* using pure protein. In a next step both the N-terminal domains of vvEDS1 and vvPAD4 (if the domain on itself can be purified separately) should be tested further for catalytic activity, including hydrolysation but also alternative reactions catalysed by α/β hydrolase fold containing enzymes, especially hydroxynitril-lyase activity.

Structure determination of PAD4 orthologues with different insertion lengths using x-ray crystallography or Cryo-EM should be done to uncover the role of the insertion and its influence on substrate binding/catalytic activity. Further *in planta* experiments based on the findings of the role of the catalytic domain in the defence against the green peach aphids should be performed. Especially further studies of different mutants influencing the active site based on the atPAD4 model and structures of known α/β-hydrolase-fold containing enzymes like KAI2, DWARF14 or GID1.

4.3. Protein-protein interactions

4.3.1. atEDS1-nanobody interactions

4.3.1.1. Mapping of the nanobody epitopes

Nanobodies binding to twenty-two different epitopes resulted of the immunisation of a llama with recombinant atEDS1 (Table 13). Of these 22 groups, 19 were found to the N-terminal domain of atEDS1 while three did not form a stable complex on a gelfiltration column with the full-length protein (figure 9). The non-binding nanobodies might bind not strong enough to form a complex stable during gelfiltration, although typically the affinity of a nanobody to its epitope is in the sub-nanomolar range, indicating a very strong interaction (Zarvtanik et al. 2018). Another possibility is that these nanobodies bind to an impurity contained in the protein provided to immunize the llama. Much more interesting is the fact that all nineteen nanobodies with affinity to atEDS1, bind to the N-terminal lipase-like domain (figure 9). While it cannot be excluded that this is simply a coincidence, this observation requires discussion. *Cameloid* single antibodies are known to prefer rigid epitopes rich in aromatic amino acids in contrast to common IgG (Zarvatnik et al. 2018). Several online tools are available to calculate the flexibility and rigidity of protein sequences. Figure 30 shows the output of three of those tools. The MEDUSA webserver (Meersche et al. 2021) predicts the flexibility of amino acids based on known x-ray structures; the X-tal pred server (Slabinski et al. 2007) combines several tools to predict the probability of a protein forming crystals based on the sequence, it includes a tool to predict disordered regions (Ward et al 2004). Finally, the PredictProtein server (Bernhofer et al. 2021) also calculates several tools to determine properties of proteins, including a prediction of the B-factor of each amino acid. Additionally, a representation of the B-factors for the atEDS1-1AT73 structure is shown in a heat map.

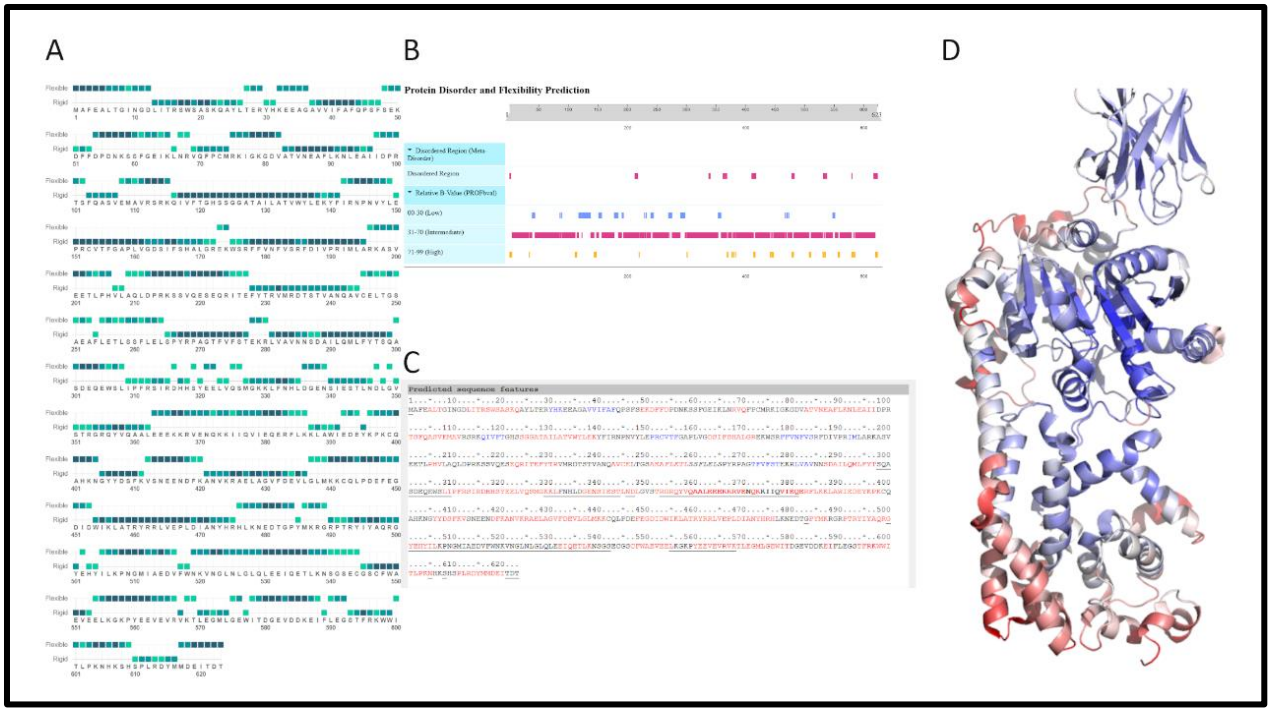

Figure 30: Analysis of the rigidity of atEDS1. Nanobodies have been shown to prefer rigid epitopes in contrast to common IgG. The analysis of atEDS1 concurs with the finding that atEDS1 specific nanobodies prefer to bind the N-terminal domain. (A) Output of the MEDUSA server that classifies single amino acids as either rigid or flexible. (B) Part of the output of the ProteinPredict tool that calculates the B-value of protein sequences. High B-values, indicating flexibility, were more frequently identified in the C-terminal domain. (C) Part of the output of the XtalPred server. Underlined amino acids are predicted to be disordered. While the structures of atEDS1 solved in this thesis disprove this, these marked patches might be more flexible than others. (D) Experimentally determined B-values of the atEDS1-1AT73 structure solved in this thesis in a heat map. Red patches indicate a high B-factor while blue patches indicate a low one.

Figure 30A shows the output of the MEDUSA server that predicts large patches of flexible amino acids for the C-terminal domain of atEDS1, although parts of the αP, αQ and αR helix are predicted to be rigid. The determination of the B-factors from the sequence by the ProteinPredict tool is shown in Figure 30B and also predicts more areas of high B-values and disordered patches for the C-terminal domain. This finding is supported by the output of the XtalPred platform that predicts disordered patches in the C-terminal domain of atEDS1 (figure 30C). While the results of the crystallographic experiments show an ordered fold for the whole protein, this can be a hint of a higher degree of flexibility in the underlined large regions of the atEDS1 EP domain. Finally figure 30D shows the B-values found for the atEDS1-1AT73 structure, that were found to be tendentially higher at the C-terminal domain compared to the lipase like N-terminal domain. Taken together, these findings suggests that indeed the N-terminal domain provides more suitable epitopes compared to the C-terminal domain, fitting to the fact no nanobody was found to bind the C-terminal domain (figure 9).

4.3.1.2. Analysis of the nanobody binding sites found in the crystal structures.

The crystal structures of three different atEDS1-nanobody complexes were solved in this thesis. An overview of the three structure is presented in figure 31.

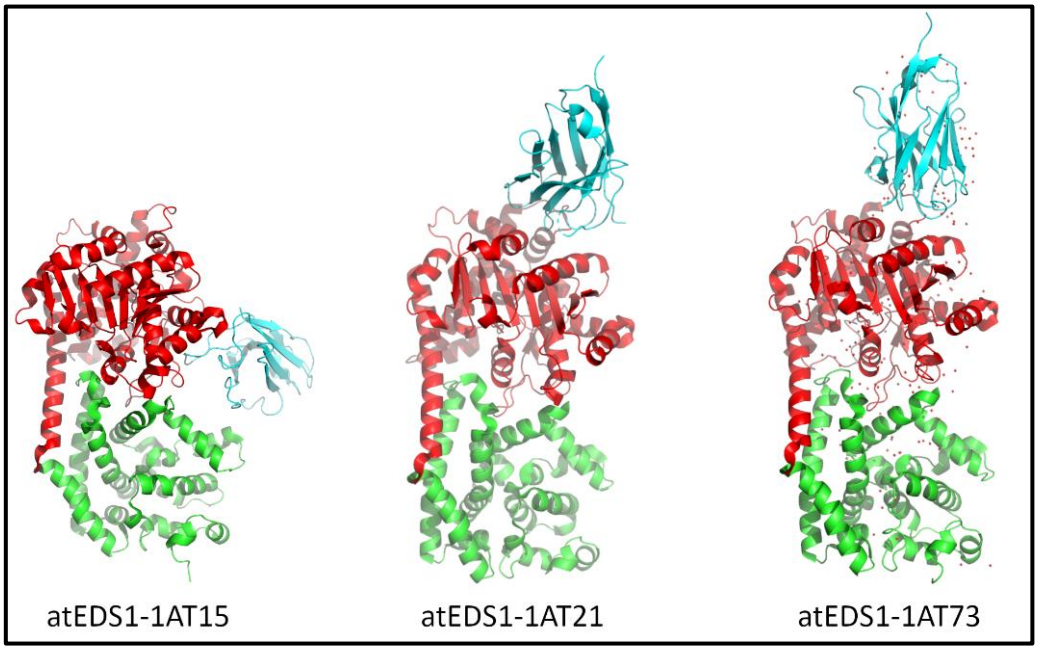

Figure 31: Cartoon representation of the three atEDS1-nanobody complex structures. The nanobody is represented in turquoise, the N-terminal lipase-like domain of atEDS1 in red and the C-terminal EP domain in green.

A more detailed analysis of the binding sites shows that, 1AT15 CDR1 binds mainly to the αB1-helix of the lipase-like domain of EDS1. Parts of the nanobody backbone binds to the αS`-helix of the EP domain in a non-CDR contact although the contact area is small compared to the other two CDRs. CDR3 binds to the region between beta-sheets one and two (figure 32).

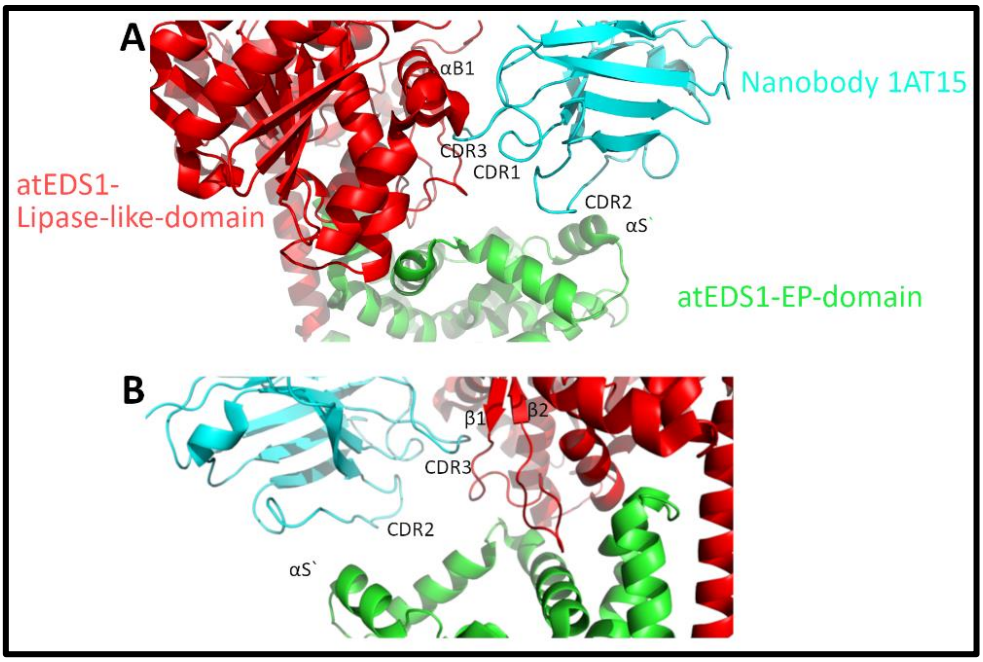

Figure 32: Detailed view of the interaction site between atEDS1 and the nanobody 1AT15. All proteins are represented in the cartoon style. The nanobody is coloured turquoise, the EP domain of atEDS1 green and the lipase-like domain red. CDR= complementary determining region of the nanobody. 1AT15 binds to the lipase like domain of atEDS1 although some interactions between the backbone following CDR2 and the helix αS from the EP domain were found. Figure A and B show the same representation turned by roughly 180 degrees.

1AT21-CDR1 binds to the region between the αF and αG-helix, while CDR2 binds to the αK helix. CDR3 of 1AT21 was positioned near the αA-helix although large parts of the loop could not be placed in the electron density.

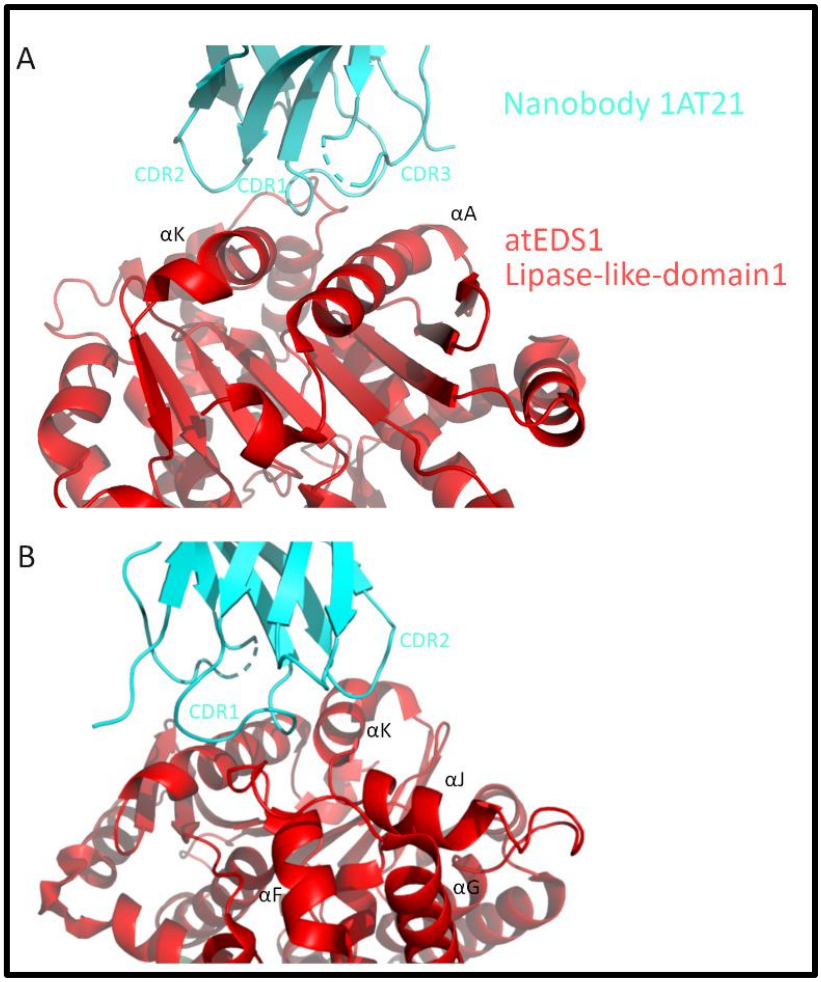

Figure 33: Detailed view of the interaction site between atEDS1 and the nanobody 1AT21. All proteins are represented in the cartoon style. The nanobody is coloured turquoise and the lipase-like domain of atEDS1 red. CDR= complementary determining region of the nanobody. 1AT21 binds to the N-terminal lipase like domain, mainly the nanobody binds between the helices αF and αG as well as the helix αK. CDR3 could not be placed in the electron density and is therefore represented in a dotted line. Figure A and B show the same representation tilted by roughly 180 degrees.

The binding site of 1AT21 and 1AT73 largely overlap, meaning the nanobodies bind to the same position although in slightly different orientations. For 1AT73 CDR1 binds to the αA-helix and both CDR2 and CDR3 bind to the region between the αF and αG helix. CDR3 also contacts the αK helix of the lipase like domain.

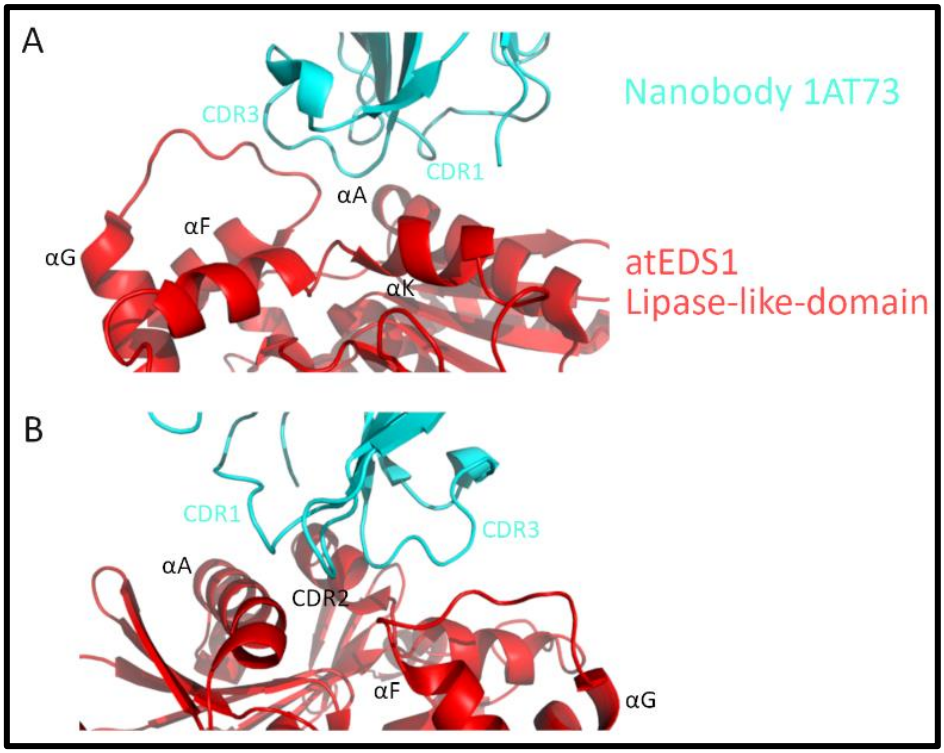

Figure 34: Detailed view of the interaction site between atEDS1 and the nanobody 1AT73. All proteins are represented in the cartoon style. The nanobody is coloured turquoise and the lipase-like domain of atEDS1 red. CDR= complementary determining region of the nanobody. 1AT73 binds to the N-terminal lipase like domain, mainly the nanobody binds between the helices αF and αG as well as the helix αK and αA. Figure A and B show the same representation turned by roughly 180 degrees.

The complex structures verify the finding that the nanobodies bind to the N-terminal domain (figure 32-34). Notably the binding sites for the nanobodies 1AT21 and 1AT73 are almost identical as shown above. Despite this finding the complexes crystallized in two different space groups with P321 (1AT21) and C2 (1AT73) and therefore two different lattice systems.

4.3.2. Does EDS1 adapt its conformation upon docking to its signalling partners SAG101 and PAD4?

One of the main questions of this thesis was to answer, if atEDS1 undergoes structural adaptations upon binding of atSAG101 or atPAD4, meaning if atSAG101/atPAD4-unbound atEDS1 does have a different conformation than the atEDS1 part of the atEDS1-atSAG101 structure (PDB: 4NFU). To that end, the individual atEDS1-nanobody structures were overlayed with the atEDS1-portion of the above structure. For the atEDS1-1AT15 and the atEDS1-1AT21 structures, no difference was observed at all regarding the EDS1-SAG101 structure. Both nanobodies bind to a different site (3.1.8.) at the N-terminal domain and the complexes crystallized in a different space group (Table 15 and 16). This indicates that atEDS1 is a rather rigid protein.

No analysis of the individual amino acids was done, as the resolution is too low to make significant observations. In contrast, the atEDS1 structure in complex with 1AT73 shows some differences compared to the atSAG101 bound state. The loops connecting the αF and αG-helix, the αC helix and β-sheet 6, as well as the loop connecting the αI and αJ-helix were placed in a slightly different conformation compared to the complex structure (Figure 35). In summary, these structural adaptations are marginal and do not suggest any functional importance.

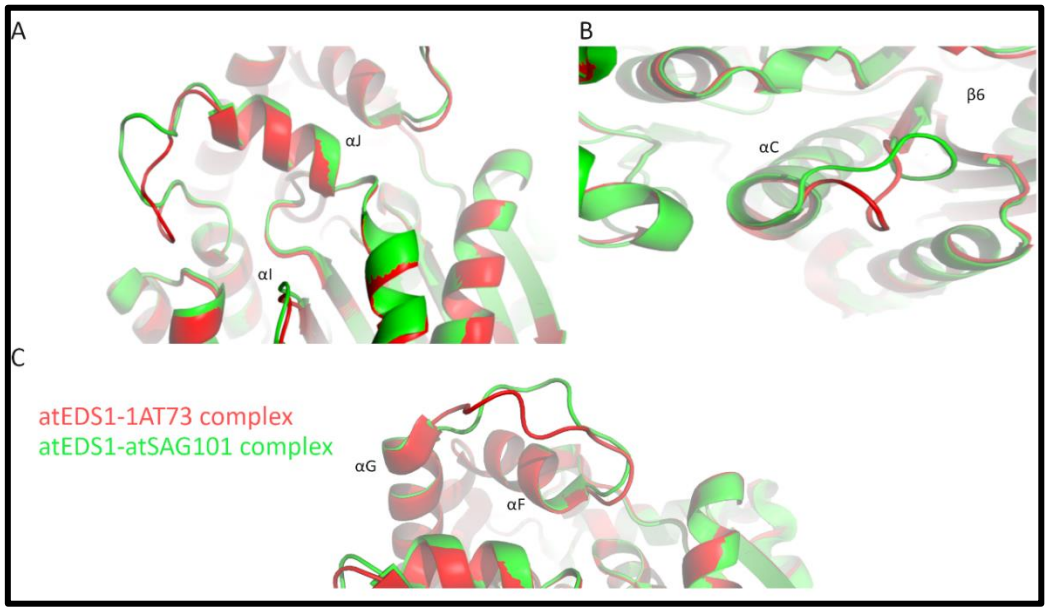

Figure 35: Overlay of the atEDS1-atSAG101 structure (green) and the atEDS1-1AT73 structure (red) without the nanobody. Several loops show a slightly different conformation when atEDS1 binds to 1AT73 in contrast to the atSAG101 bound state. Figure A shows the loop between αI and αJ. Figure B shows the loop between αC and β6. Figure C shows the loop between αG and αF.

The αH helix is of special interest since it is important for the formation of the atEDS1-atSAG101 complex (Wagner et al. 2013). In both the atEDS1-1AT15 and atEDS1-1AT21 structure, the helix is formed identical to the atEDS1-atSAG101 complex structure. In both cases the helix was also involved in a crystal contact to another copy of atEDS1. In the atEDS1-1AT73 structure the αH helix is not structurally established at all and was instead found to be a molten, unstructured loop, indicating the flexibility of the region when not bound to another protein. The helix αH is essential for the interaction of atEDS1 to both atSAG101 and atPAD4 (Wagner et al. 2013) and mutation at key points of the helix can implicate the defence against certain pathogens (Wagner et al. 2013). The helix seems to be in a molten, disordered state until binding to either SAG101 or PAD4. The effector protein HopA1 is also postulated to bind the αH helix and might therefore disrupt or prevent the formation of atEDS1-atSAG101/atPAD4 complexes (Park et al. 2015).

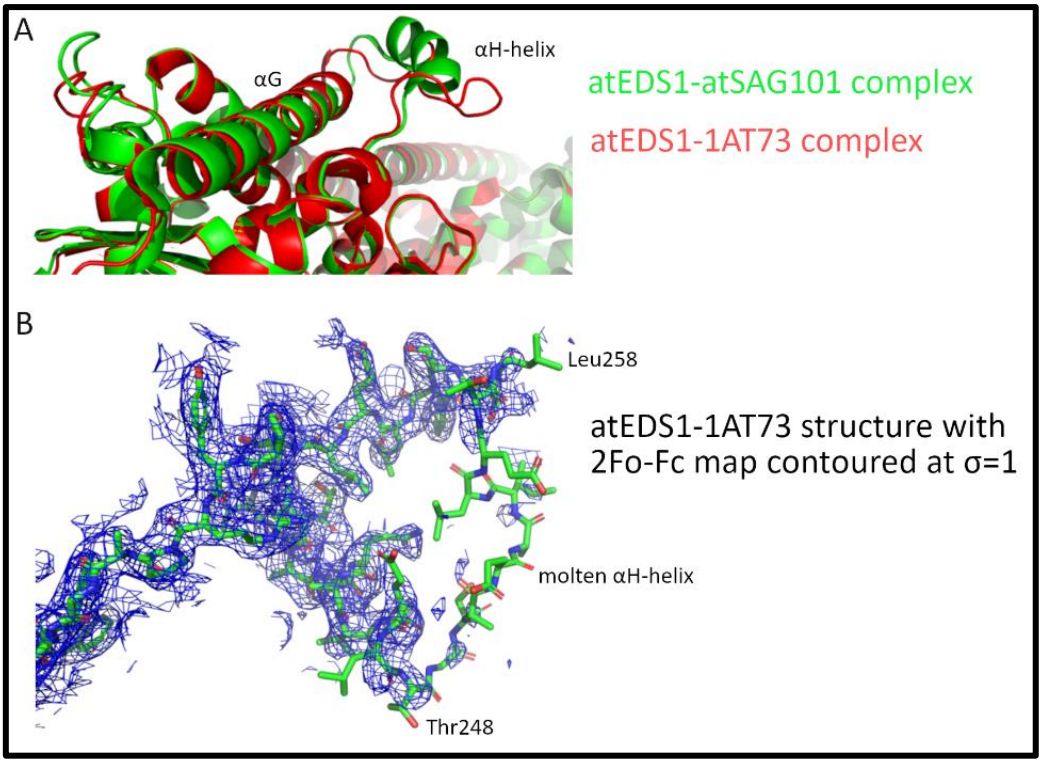

Figure 36: Comparison of the αH helix of atEDS1 either bound to atSAG101 or 1AT73. Figure A shows an overlay of both structures without binding partners in the cartoon representation. The helix that is essential for the interaction with atSAG101 is not formed in the 1AT73 bound state but was placed as a molten, disordered loop. Figure B shows the electron density map (2Fo-Fc) in the 1AT73 bound state at a contour level of σ=1. No density was observed for the αH loop between Thr248 and Leu258.

Taken together, atEDS1 and combined with the SAXS-analysis done using His tagged atEDS1 (Voss et al. 2019) it can be concluded that atEDS1 is a rather rigid protein that does not undergo large conformational changes upon binding to atSAG101. A remarkable exception is the helix αH: it is molten in the atEDS1-1AT73 structure, but well established in the atEDS1-atSAG101 complex (Wagner et al., 2013). This interesting observation emphasizes the critical relevance of this helix as interaction platform for EDS1's signalling partners atSAG101 and atPAD4.

4.3.3. The puzzle of the oligomeric state of EDS1

As described in the introduction, several studies were published with evidence of atEDS1 self-interaction (see 1.3.1.). Additionally, atEDS1 is noted to form homodimers in the UNIPROT database[13].

Contrary to this, atEDS1 recombinantly expressed in *E.coli* never formed complexes in any of the gelfiltration experiments performed by me, Christine Tölzer and Stephan Wagner. The discrepancy of this finding to the current state of the literature is striking; in the following paragraphs possible explanations are provided.

4.3.3.1. atEDS1 can form extensive self interactions in crystal packings

The first attempt to explain the studies claiming a homo-dimerization of atEDS1 is to analyse the interaction found in the atEDS1-nanobody crystals. The very nature of crystal formation requires interaction between individual atEDS1-molecules and/or the nanobodies. To distinguish between a biological relevant protein-protein contact or a solely crystallographic contact, software can be used to determine the area of the contact site and the involved energy. The three structures were analysed using the PISA (Krissinel & Henrick 2017) and EPICC servers (Duarte et al. 2012). Based on the contact area and energy, the contacts are then judged to be either biologically relevant or not. Due to the low resolution, I did not analyse the binding energies of the 1AT15 and 1AT21 containing structures.

In general, all three structures were evaluated by both servers to be likely monomeric in solution, in line with the thesis of a monomeric atEDS1. Nevertheless, the protein contacts in the different crystals were further analysed as weak interactions might explain the positive assays discussed above. All mentioned values below were calculated using both servers, the calculated interaction area deviated of 5 $Å^2$ or less between both servers.

The analysis begins with the atEDS1/nanobody interfaces since they can serve as prime examples for weak and non-obligatory protein/protein interactions that are nevertheless strong enough to provide stable complexes in solution.

In the case of the atEDS1-1AT15 structure, the servers reported an atEDS1/nanobody interface area of 775 $Å^2$. For the 1AT21-atEDS1 binding site an interaction area of 694 $Å^2$ was calculated and finally the analysis of the atEDS1-1AT73 interaction area resulted in a calculated interaction area of 768 $Å^2$. These values fit the published expectation of nanobody-epitope binding sites of 750 ± 180 $Å^2$ (Zavrtanik et al. 2018).

13 https://www.uniprot.org/uniprot/Q9SU72#interaction

For the atEDS1-1AT15 crystal, the server identified several interactions between neighbouring atEDS1 molecules that were all classified crystallographic contacts. Notably, two interactions with neighbouring atEDS1 molecules have interfaces sizes of 1185 $Å^2$ and 855 $Å^2$, respectively, i.e., they are larger than the atEDS1-nanobody interaction sites and candidates for transient atEDS1-self interactions in solution. Regarding the atEDS1-1AT21 crystal three interaction sites between the copies of atEDS1 were identified involving an area of 957 $Å^2$, 770 $Å^2$ and 143 $Å^{2.}$ Again, two of the atEDS1-atEDS1 interaction exceed the area involved in the nanobody-atEDS1 binding site.

Notably, for both the atEDS1-1AT15 and the atEDS1-1AT21 structure EPPIC proposed a trimeric assembly (figure 37) with a certain probability requiring only the atEDS1 molecules. Both assemblies were scored with a probability of 10% to be biologically relevant. For these assemblies, the used contact side involves 860 $Å^2$ for the atEDS1-1AT15 and 957 $Å^2$ for the atEDS1-1AT21 structure. In both assemblies, the atEDS1 protomers are arranged in a very similar way despite the vastly different nanobody binding site and crystallization condition, lending validity to a potential biological significance. Notably the helices αH, αG and αP that are also essential for the atEDS1-atSAG101 interaction, are heavily involved in forming the trimeric assembly shown in figure 37.

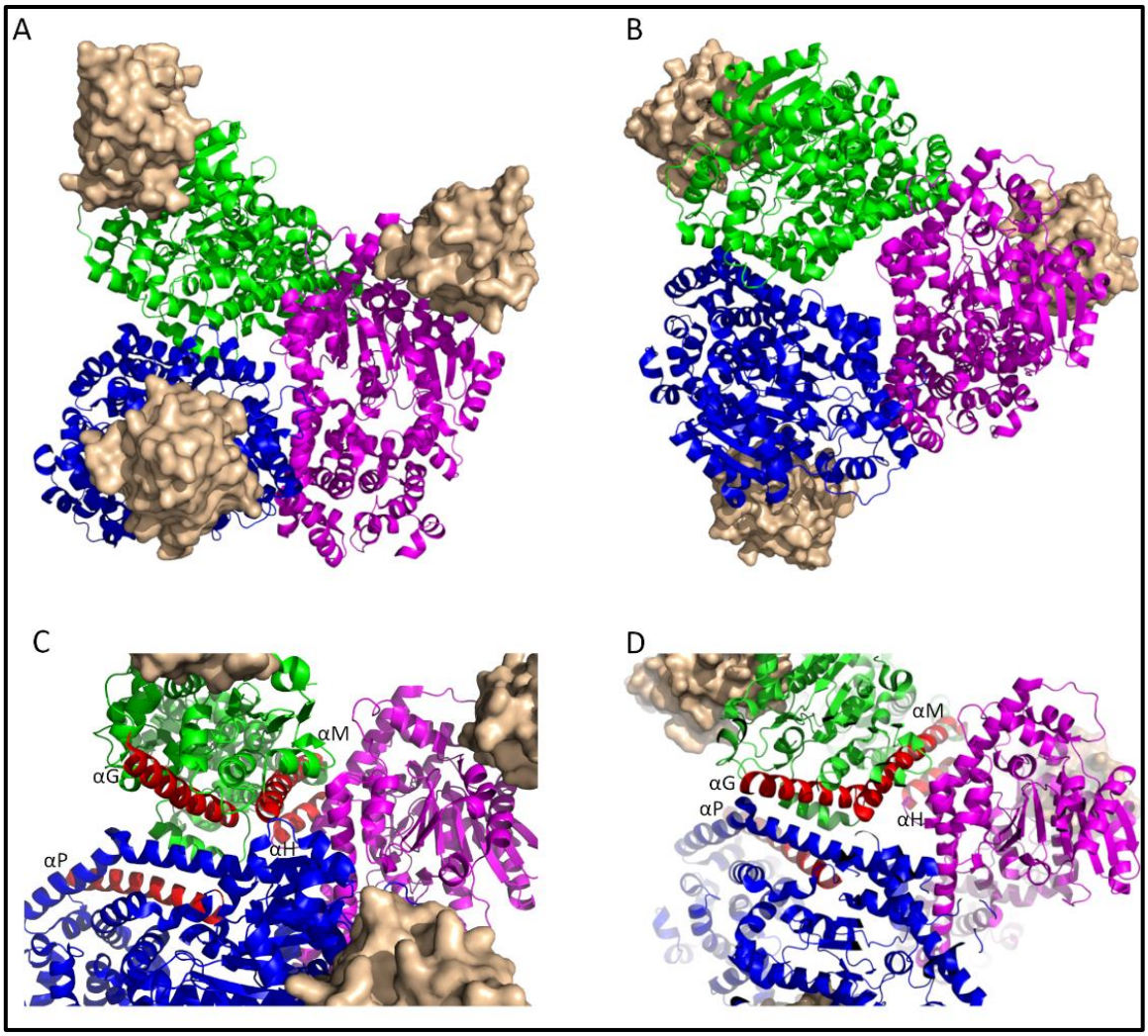

Figure 37: Detailed view of crystallographic protein-protein contacts of the atEDS1-1AT15 (B and D) and atEDS1-1AT21 (A and C) crystal structures. Protein-protein interactions were analysed using the PISA and EPPIC servers, all contacts were determined to be crystallographic and likely not biologically relevant. Both structures share common interaction sites that lead to the shown trimeric conformation. The interaction heavily involves the helices αM as well as αG, αH and αP, the latter being otherwise responsible for the interaction with atSAG101 or atPAD4. Nanobodies are shown as beige surface representation, the atEDS1 molecules are depicted in the cartoon representation in magenta, blue and green. Key components for the protein-protein interaction are highlighted in red in panels C and D. Figure A shows an overview of the trimer found in the atEDS1-1AT21 structure. Figure C shows the same ensemble focused on the atEDS1 molecules with the key elements for the interaction highlighted in one copy of atEDS1 each. Panel B and D show the same representations for the atEDS1-1AT15 structure.

While this trimeric assembly is not found in the atEDS1-1AT73 containing structure, a second potentially relevant interaction occurs solely in this crystal. It involves a head to tail arrangement in a manner that might lead to linear filament growth as shown in figure 38. The interaction involves the αG and αI and αJ helices of the N-terminal domain as well as the αM, αN and αP helices of the EP-domain. This might correlate to the tendency of atEDS1 to aggregate with time (figure 38).

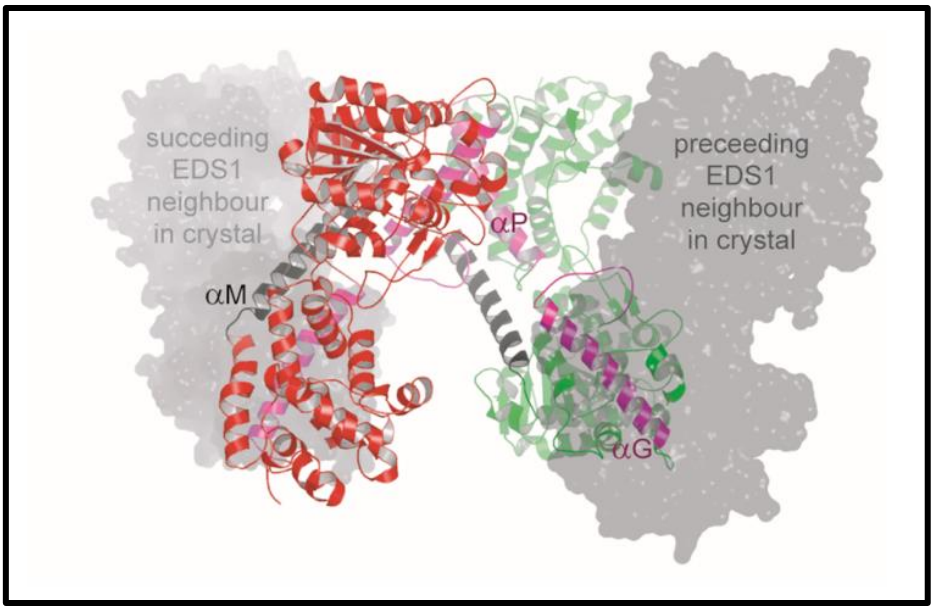

Figure 38: Crystallographic protein-protein interactions occurring in the atEDS1-1AT73 crystal-structure. A series of interactions heavily involving the helices αM, αP and αG was found to form a potential linear aggregate. All interactions were classified as purely crystallographic by the PISA and EPPIC servers. Figure taken from Voss et al. (2019).

In conclusion, the analysis of the crystal packing supports the thesis of a monomeric unbound atEDS1. Nevertheless, a trimeric assembly involving a considerable interaction area was identified in addition to an assembly leading to a linear accumulation of atEDS1 molecules. These interactions are weak and transient, meaning they do not lead to stable atEDS1 oligomers in solution; nevertheless, could provide a clue to explain the positive Y2H results identified in several publications (Feys et al. 2001, Wagner et al. 2013, Bhattacharjee et al. 2011).

4.3.3.2. Oxidative disulphide bond formation: the key for atEDS1's self-interaction propensity?

Another avenue to rationalize the atEDS1-atEDS1 interactions identified via Y2H lies in the tendency of atEDS1 to form aggregates. Solutions of purified atEDS1 show visible aggregation after 2-3 days storage at 4°C. The analysis of the dependency on reductive agents as well as the possible involvement of several cysteines has been discussed in chapter 4.1.1., nevertheless, this point values a mention here, as the formation of covalent atEDS1 oligomers due to the formation of disulphide bridges is a valid theory to explain some of the studies suggesting an atEDS1 self-interaction. The analysis of the oligomeric state of atEDS1 over time if fresh DTT is added daily is shown in figure 13.

4.3.3.3. Non-crystallographic evidence that atEDS1 is a monomer in its atSAG101/atPAD4 unbound form

Several additional factors must be taken into consideration before the final determination of the oligomeric state of atEDS1. Due to the problems of generating protein crystals of atEDS1 with a good diffraction quality, SAXS-studies were performed by Christine Toelzer before the generation of the nanobodies, that were successfully used as crystallization chaperones in this thesis.

The small angle X-ray scattering of a protein in solution is suitable to provide a low-resolution structure only if the protein is largely monodisperse; therefore, a size exclusion chromatography-SAXS tandem experiment was performed in which the purified atEDS1 sample was applied to a gelfiltration column in order to separate all aggregated protein particles. The low-resolution model gained (Voss et al. 2019) confirms the finding of the gel-filtration experiments and the analysis of the protein interactions in the crystal as a monomer of atEDS1 fits nicely in the calculated density.

The disadvantage of using eukaryotic proteins recombinantly expressed in bacteria is the fact that most post-translational modifications that might exist in the natural environment are missing from the protein. It is a valid argument to claim the interactions found *in planta* and in the Y2H assays might rely on modifications that are not done by *E.coli*. This point, however, is disproven by the Co-IP experiments shown in figure 39.

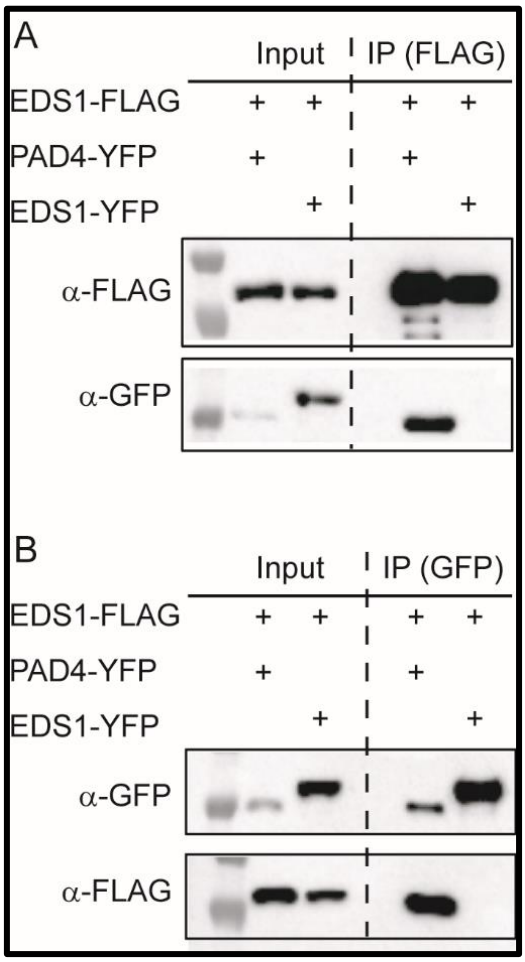

Figure 39: In planta immunoprecipitation assays of atEDS1 and atPAD4. Both proteins were recombinantly expressed in Nicotiana tabacum with either a FLAG- or a YFP-tag. Cell lysates were gathered and immunoprecipitation experiments with immobilized antibodies recognizing either the FLAG tag (A) or the YFP tag (B) were performed. Subsequently, proteins were detected via western blotting. The results show that atEDS1 does bind atPAD4 but not itself. Figure is taken from Voss et al. 2019.

AtEDS1 was transiently expressed in *N.benthamiana* with either a FLAG or yellow fluorescence protein (YFP) tag, atPAD4 with a YFP tag was used as a positive control, as the interaction of atEDS1 and atPAD4 is well known. After immunoprecipitation using either anti-FLAG or anti-YFP antibodies no atEDS1 with the opposite tag was found. This indicates that, even expressed in a plant atEDS1 does not interact with itself. Finally, the gelfiltration experiments done with *Arabidopsis* leaf-extract by Feys et al. 2005 should be taken into consideration (figure 40).

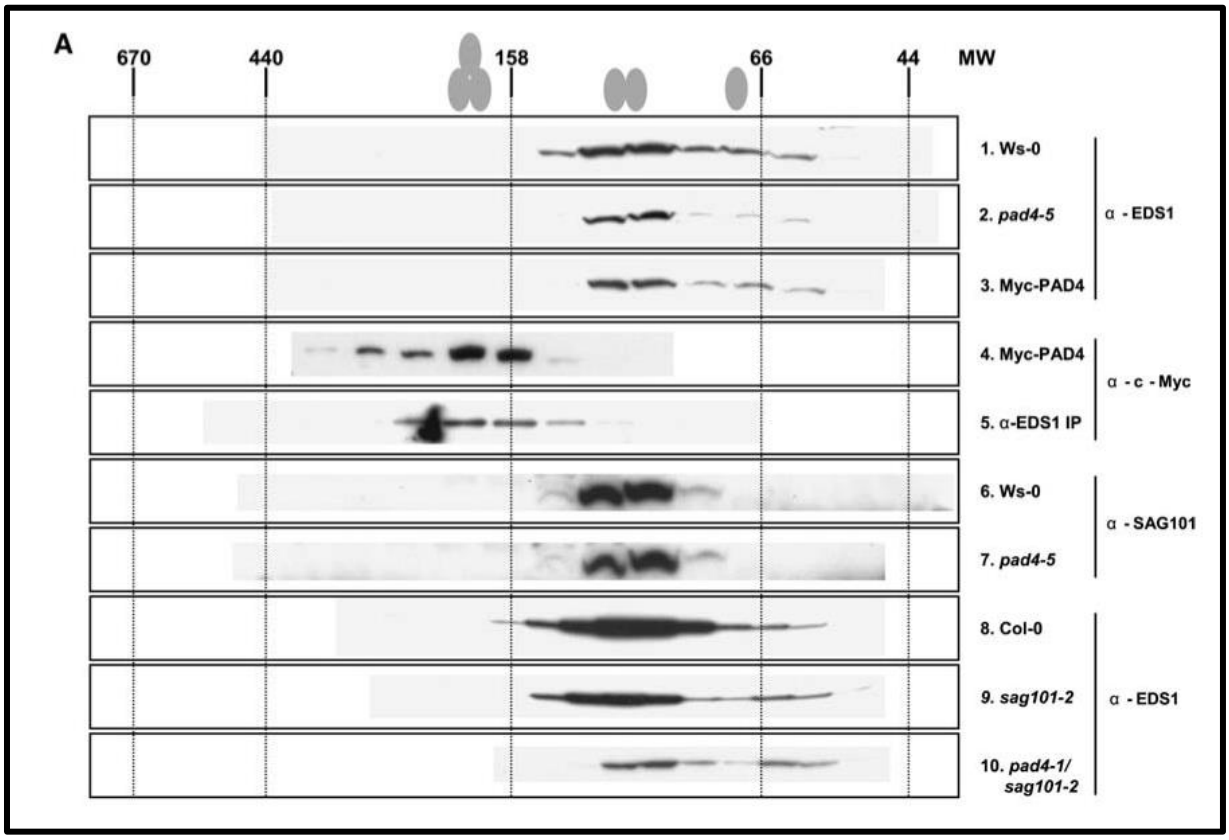

Figure 40: Soluble leave extract from A.thaliana was applied to a gelfiltration column. Afterwards a western-blot analysis to detect atEDS1, atSAG101 and atPAD4 containing complexes was performed. The used antibodies are shown on the right side, the molecular weight determined by calibration with known proteins is shown above, with notions where a monomer, dimer or trimer of atEDS1/atPAD4/atSAG101 is expected to occur. The genotype of the plants is shown next to the lanes. Ws-0 and Col-0 are different A.thaliana variants. Pad4-1, pad4-5, and sag101-2 are lines that do not possess the gene in question. All experiments show a pool of monomeric atEDS1 but no monomeric atSAG101 or atPAD4. Taken from Feys et al. 2005, with kind permission of Oxford academic press.

The experiments done with different *A.thaliana* variants as well as several lines without atPAD4 and/or atSAG101 all show a pool of monomeric atEDS1 but no monomeric atPAD4 or atSAG101. This finding validates the hypothesis that atEDS1 is a monomer when unbound to the latter. It can be concluded that the plant cells have a pool of atEDS1 that is then available to bind to either SAG101 or PAD4. As of now, no data are published that show a function of atEDS1 in absence of both binding partners atPAD4 and atSAG101 in contrast to atPAD4 that was shown to act independently in defence against the green peach aphid (Dongus et al. 2020).

Nevertheless, the studies reporting a self-interaction of atEDS1 need to be discussed as well. The observed co-immunoprecipitation of atEDS1 by atEDS1 (Bhattacharjee et al. 2011) can be explained due to indirect, bridging interaction via larger complexes. Additionally positive BiFc experiments (Bhattacharjee et al. 2011) might either be explained by false positive results, a distinct possibility for these experiments (Horstman et al. 2014), or due to a physical closeness of two atEDS1 copies as a result of a larger complex.

Nevertheless, several studies showed a self-interaction of atEDS1 in Y2H experiments (Wagner et al. 2013, Feys et al. 2001), giving credibility to the thesis of a homo-dimerization.

The findings described above do not contradict that several copies of atEDS1 might be part of a larger protein complex involved in the immune response. Recently, several large complexes involved in the plant-immune response have been either described or solved. These large complexes include the CC-NB-LRR receptor ZAR1 containing complex (Wang et al. 2019) that is not dependent on the EDS1 family, that forms a calcium channel through the cell membrane upon activation and finally leads to cell death (Bi et al. 2021) as well as the TIR-NB-LRR receptor RPP1-tetramer that shows NADase activity and is dependent on the EDS1 family (Ma et al. 2020). AtEDS1 itself has been shown to interact with the CC-NB-LRR protein NRG1 that is a hypothesized to form oligomers (Sun et al. 2021).

Very recently, atEDS1 and atPAD4 have also been shown to form a complex with the CC-NB-LRR protein ADR1 and the membrane bound receptor kinases BAK1 and SOBIR1 as well as the cytoplasmatic kinase PBL31 (Pruit et al. 2021). These studies that all involve multiprotein assemblies show that the plant immune system relies on a series of complexes that may very well contain several copies of the EDS1-family, therefore easily explaining the published co-localization results despite the findings of this thesis. In conclusion, a model comprising all existing evidence on atEDS1 structure and function can be developed (Figure 41): atEDS1 is a monomer in its basic state but is a part of a complex network of proteins that activates and regulates main pathways of the plant immune response.

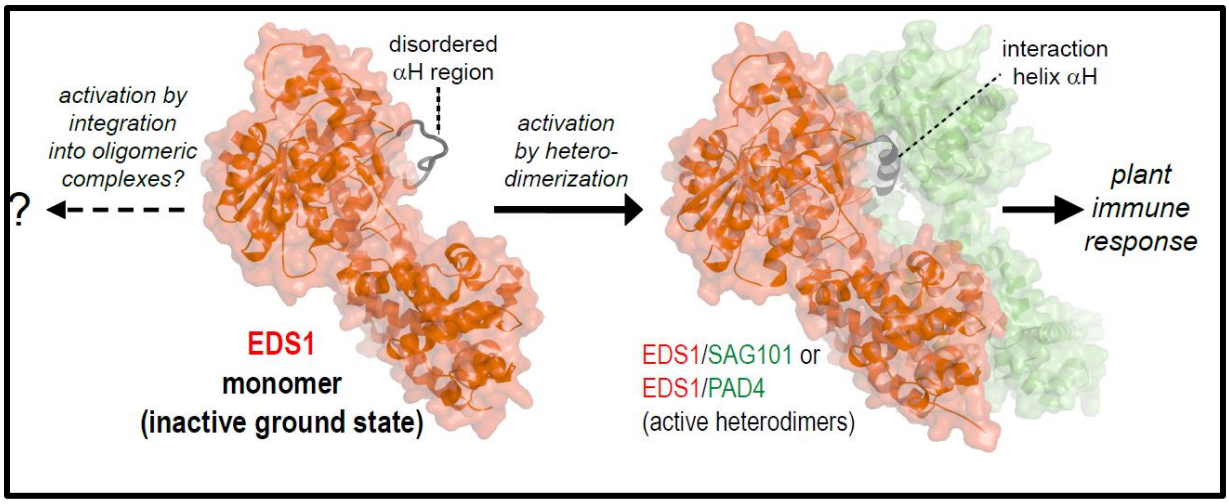

Figure 41:Model derived from the experiments done in this thesis. AtEDS1 is a monomer in its ground state. The αH region is disordered when EDS1 is in its ground state. Upon binding to either PAD4 or SAG101 the αH helix is formed and becomes part f the interaction site. We propose that EDS1 is inactive in its monomeric state, while its heterodimers are essential to trigger the immune response.

4.3.4. Analysis of the vvEDS1^{Nterm} dimer structure

4.3.4.1. The vvEDS1^{Nterm}protein is a consequence of intron retention.

Before analysing the protein structure, itself, first the sequence leading to this truncated protein needs to be discussed. As shown in figure 22 the protein found during the expression of vvEDS1 was not the expected full-length ~70kDa vvEDS1, but a truncated form consisting of the major part of the N-terminal domain with just the helices αL and αM missing. Upon sequencing, it was found that this is the result of an alternative splicing event, in concrete the retention of the second intron (figure 23A). In contrast, the first intron was removed correctly during splicing (figure 23B). Hence, the question arises if this unexpected find is the consequence of a biological relevant alternative splicing event or an experimental artifact. The sequence used was cloned from a cDNA library of *Vitis viniferia-cabernet sauvignon* inoculated with powdery mildew. At the time the thesis was written, an attempt to verify the alternative splicing event occurs *in planta* by creating a new cDNA library was still in process. Thus, the verification of a potential biological relevance is still pending.

The first hint that this protein variant is not an artifact but biological relevant is the position of the new stop codon itself. Only seven amino acids are exchanged at the N-terminus of the new protein (figure 23C), additionally the new stop codon is located at the border between the two domains. The last structural element found in the vvEDS1^{Nterm} structure was the strand β10 with the helix αL potentially made up by the remaining N-terminal amino acids not defined in the electron density. Thus, only the helix αM, that is needed to connect both domains in atEDS1 is missing in the truncated vvEDS1^{Nterm}.

Combined with the recent find that the N-terminal lipase like domain of the EDS1-protein family member PAD4 from *A.thaliana* is sufficient for defence against the green peach aphid (Dongus et al. 2020), this is a potential hint that the resulting protein is a valid isoform of vvEDS1 as all elements of an α/β-hydrolase fold are present.

Nevertheless, until now no role of an isolated EDS1 lipase-like domain has been published and no studies showing the presence of such a protein are known. As such, the potential use and validity of such an alternative splice variant will be discussed at this point. In general, it has been found that 70% of all multi-exon genes undergo alternative splicing (Chaudhary et al. 2019). While the predominant mode of alternative splicing in human is exon-skipping, roughly 60% of the occurring alternative splicing elements in plants are made up of intron retention and leads mostly to premature stop codons (Chaudhary et al. 2019). Thus, the splicing event identified for vvEDS1 is neither unexpected nor a rarity but rather the norm for plants. Premature stop codons can lead to nonsense mediated decay (NMD), a process that leads to the removal of the 5`cap found on mRNAs and subsequently to degradation of the mRNA (Lykke-Andersen & Jensen 2015). The degradation of mRNAs via the NMD pathway is triggered either by a large 3´-untranslated region or the presence of an exon-junction complex after the stop codon. Exon-junction complexes are formed at the edges of the exons that are fused together after splicing, these complexes are then removed by the ribosome during translation. If a stop codon is located downstream of an exon-junction complex, the degradation of the mRNA is triggered. (Lykke-Andersen & Jensen 2015). While this is not the case for vvEDS1, as the exon junction complex formed at the splice site removing the first intron is removed during the first translation, NMD can alternatively be triggered by a long 3´untranslated region (Lykke-Andersen & Jensen 2015). The length of this region was specified to be larger than 350 nucleotides (Kalyna et al. 2012). As the last 234 amino acids are cut off due to the new stop codon in the vvEDS1 transcript (figure 23C) this condition is met, as at least 702 (234*3) nucleotides are not translated. Therefore, the intron retention leading to the truncated vvEDS1 might lead to degradation of the mRNA and might play a role in the regulation of the immune response. At the same time, many transcripts containing premature stop codons are resistant to nonsense-mediated decay in *Arabidopsis thaliana* (Kalyna et al. 2012). It is therefore unclear if the mRNA retaining the second intron is degraded or not *in planta* without further experimental data.

Several aspects of mRNA with retained introns need to be mentioned at this point. Partially spliced transcripts can be retained in the nucleus where they either undergo degradation or are stored and released upon external stimuli (Monteuuis et al. 2019). Interestingly, several studies have shown a connection between nonsense-mediated decay and plant immunity, although it must be noted that all following studies were performed in *Arabidopsis* and no information about splicing or nonsense mediated decay in *Vitis vinifera* is available.

Nonsense-mediated mRNA decay is downregulated upon infection with *Pseudomonas syringae* and plants unable to downregulate NMD are more susceptible to such an infection (Gloggnitzer et al. 2014). An NMD defective plant with inactive SMG7 gene has been shown to have an autoimmunity-stunted phenotype that could be rescued by knocking out either atPAD4 or atEDS1. Interestingly, a knockout of NDR1, a protein essential for CC-NB-LRR dependent immunity, did not rescue a smg7 phenotype, therefore linking NMD, EDS1 and PAD4 to TIR-NB-LRR dependent immunity (Riehs-Kiernan 2012). In addition, several studies found an elevated level of EDS1, PAD4 and/or TIR-NB-LRR proteins in NMD deficient plants (Rayson et al. 2012, Gloggnitzer et al. 2014, Jeong et al. 2011). NMD is therefore an important regulator of the *Arabidopsis* immune response. Thus, it is plausible to propose that the partially spliced mRNA of vvEDS1 is a result of the lowered NMD after infection of the wine plant with powdery mildew. The next step should be to verify the presence of this particular mRNA in a second cDNA pool and to compare the levels of the transcript against fully spliced vvEDS1 before and after infection with powdery mildew. In retrospect it must be noted that splicing is a complicated, highly regulated process that is even influenced by epigenetics as plants have been shown to have a "splicing memory" (Chaudhary et al. 2019). Additionally, cytoplasmatic splicing has been observed in conjunction with the unfolded protein response (Ricci et al. 2021).

4.3.4.2. Structural implication of the vvEDS1^{Nterm} dimer

The structure of the N-terminal domain of vvEDS1, resulting from an alternative splicing event is shown in figure 42.

Figure 42: Content of the asymmetric unit of the structure of the N-terminal domain of vvEDS1. Waters and ethylene glycols are not shown. The structure is presented in the cartoon format, helices are shown in red, β-sheets are shown in yellow and loops in green. An α/β-hydrolase fold was identified consisting of eight central beta sheets surrounded by several α-helices. The structure is very similar to the N-terminal domain of atEDS1.

As shown in figure 24B, the gelfiltration analysis of the N-terminal domain of vvEDS1 indicated a dimerization of the protein. This finding is verified by the analysis of the structure of the protein. While the asymmetric unit does only contain a single copy of the protein (table 17) an analysis using the PISA server (Krissinel & Henrick 2007) assigns a dimer as the correct biological assembly (Figure 43). This prediction is based on an interface size of 1467 $Å^2$ and an energy gain of -29.2 kcal/mol upon dimerization.

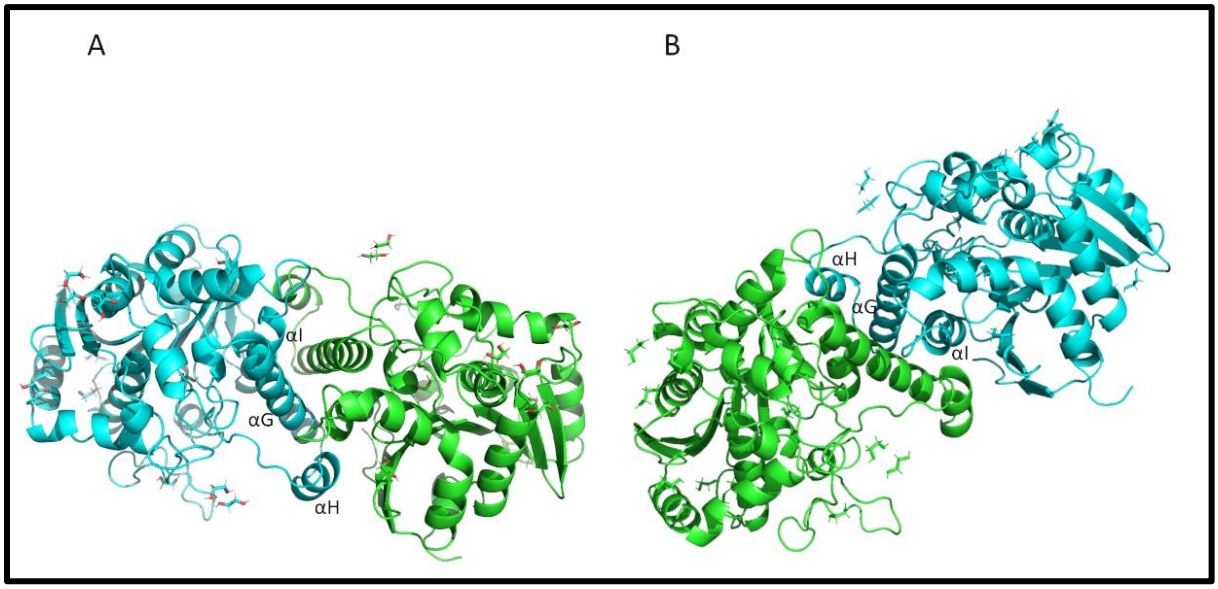

Figure 43: Biological assembly as predicted by the PISA server of the vvEDS1^{Nterm} dimer in cartoon representation. The helices αH, αG and αI that are critical for the dimerization are labelled in figure A and B for the cyan coloured variant. Figure A and B show the same representation turned by roughly 90 degrees.

As shown in figure 43 the helices αG, αH and αI are essential for the interaction and therefore are not available to interact with vvPAD4. The model for the atEDS1-atPAD4 interaction (Wagner et al. 2013) proposes that, identically to the atEDS1-atSAG101 structure, the helixes αG, αH and αI are heavily involved in the interaction of atEDS1 and atPAD4. As in the vvEDS1^{Nterm} structure, these secondary structure elements are not available for binding, this finding supports the gelfiltration experiments that show no interaction between the full length vvPAD4 and the N-terminal domain of vvEDS1 (figure 24A).

Interestingly, the dimeric structure visible in Figure 18 does not contradict a monomeric quaternary structure of SAG101/PAD4-unbound EDS1. Figure 44 shows an overlay of the N-terminal domain of vvEDS1, as shown above, with a full length vvEDS1 predicted by Phyre2 (Kelley et al. 2015). Both the helix αL that is not established in the structure of the vvEDS1^{Nterm} dimer as well as the helix αM that is cut off due to the alternative splicing, clash sterically with the vvEDS1^{Nterm}. The same is true for the potential interaction with vvPAD4 as shown in figure 44C, once again a steric clash prevents a potential interaction. This finding is in line with the gelfiltration experiments performed by me as shown in figure 24A where no interaction between the two proteins was observed. In contrast, the interaction of vvPAD4 and full length vvEDS1 has been shown via BiFc (Gao et al. 2012). To observe this interaction between both full-length proteins is an attractive goal for future studies, especially since vvPAD4 has been shown not to interact with atEDS1 while vvEDS1 interacts and functions with atPAD4 (Gao et al. 2012).

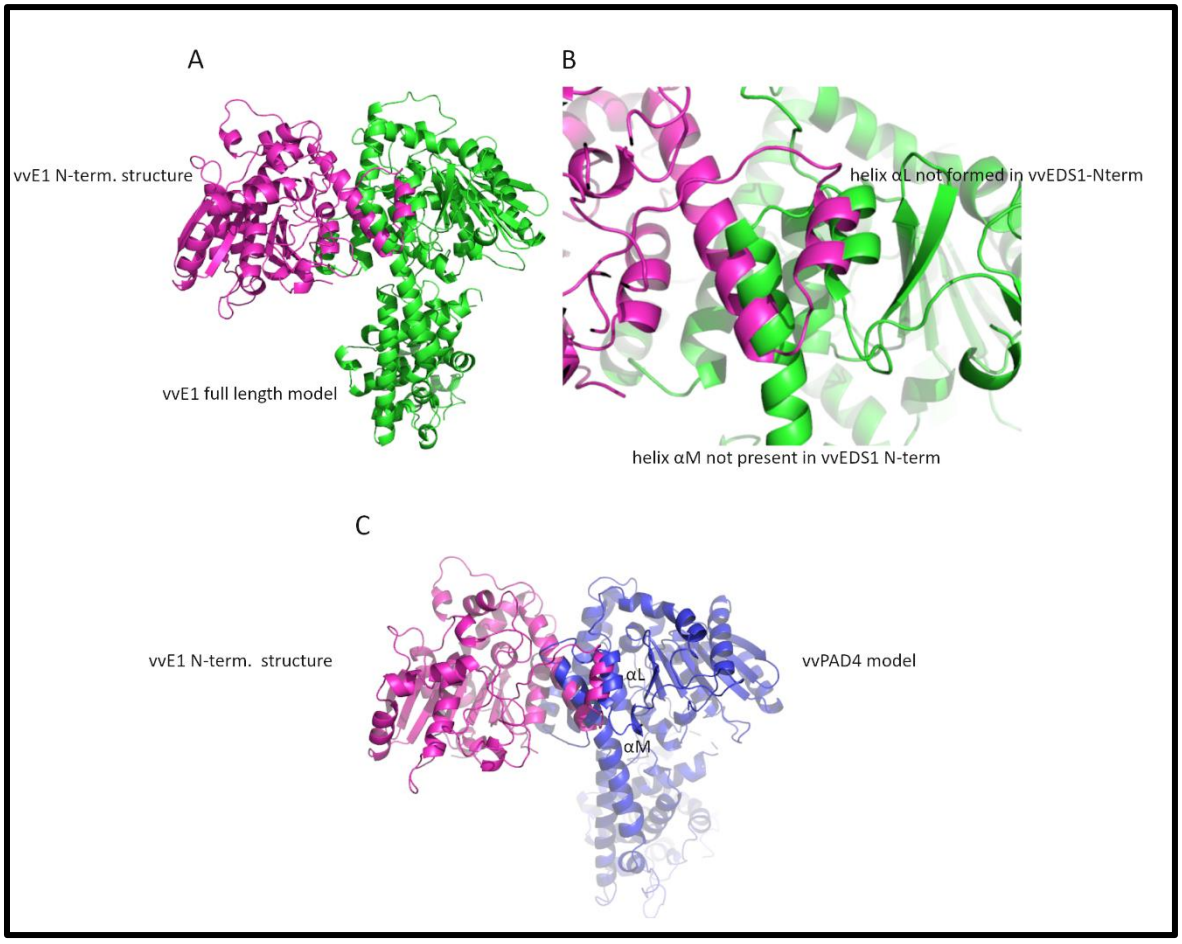

Figure 44: Hypothetical interaction of vvEDS1^{Nterm} with full length vvEDS1 (A and B) or vvPAD4 (C). The figures demonstrate that the binding as found in the vvEDS1^{Nterm} structure is not applicable to a hypothetical vvEDS1 dimer or an interaction of vvEDS1^{Nterm} with full length vvPAD4. In all cases the helices αL and αM that are either not formed or not expressed in the N-terminal vvEDS1 structure physically clash, preventing an interaction. All proteins are represented in the cartoon format, vvEDS1^{Nterm} is shown in magenta, vvEDS1 full length in green and vvPAD4 in blue. The vvPAD4 and vvEDS1 full length proteins are models build by the Phyre2 web service while the vvEDS1^{Nterm} structure was determined experimentally in this thesis.

4.4. Conclusion

The three proteins of the plant-specific EDS1 family - EDS1, PAD4 and SAG101 - are all central players that act downstream of the immune receptors. Except for SAG101, that is not present in monocots, all three proteins are conserved and found in most higher plants (Baggs et al. 2020). While the importance of these proteins is known for almost 25 years (Parker et al. 1996) the actual function of the proteins in the cell is as of yet unknown.

The challenge to feed a growing world population at times of climate change and the ease to change the genome of plants via Crispr-Cas methods (Zhu et al. 2020) raises the importance to firstly understand and secondly manipulate and improve the response of plants against stress and pathogens. The recent discovery of several EDS1 and PAD4 interacting proteins as well as the successful structure determination of several immune-receptor complexes (Wang et al. 2019, Ma et al. 2020) has significantly increased the understanding of the plant immune response, although several important questions remain unanswered. The EDS1/PAD4 and the EDS1/SAG101 complex are essential for the signalling of an activated immune receptor to efficiently trigger the immune response of the plant to pathogens of all kinds, including fungi, bacteria, virus and feeding insects. To decipher the intricacies of the plant immune response studies done *in vitro, in planta* and *in silico* are needed with scientists of all branches of biosciences working together.

This work answers some questions regarding atEDS1, but most importantly provides a necessary tool to investigate the EDS1/PAD4 complex *in vitro*. The atEDS1-nanobody structures firmly dispels reports claiming an atEDS1 homodimerization and at the same time cements the classification of atEDS1 as a pseudoenzyme with an established, but incompetent catalytic machinery. Further research using the large number of unused nanobodies that bind different epitopes of atEDS1 might lead to further insights into the behaviour of atEDS1 in its monomeric ground state. At the same time this thesis proves the advantage of the usage of nanobodies as crystallisation chaperones and establishes this as an important tool for further crystallographic experiments on EDS1 family members.

The discovery of an unexpected splice variant of vvEDS1 raises the question of the role of alternative isoforms of EDS1 before and after infection with the various pathogens. While as of now a biological relevance of this finding needs to be validated, the hypothesis of a role of the N-terminal domain of EDS1 itself in the immune response is intriguing especially as recently this was demonstrated to be valid for atPAD4 (Dongus et al. 2020).

The successful and reliable expression and purification of PAD4 from *Vitis viniferia* is an important step on the way to a EDS1/PAD4 complex structure, as well as a platform to characterise EDS1/PAD4 and its interactors *in vitro*.

As of now no successful purification of any PAD4 orthologue has been published, highlighting the value of this finding. Should an EDS1/PAD4 complex be achievable using the *Vitis viniferia* variants of PAD4 and EDS1 it will be the first time this central complex of the plant immunity is accessible *in vitro* to study.

While the characterisation and structure determination of a EDS1/PAD4 complex is not of immediate economic interest, it is an important step to understand the principles of the plant immune response as discussed above. The rational manipulation of the plant genome to create resistant or more resilient phenotypes that might increase the yield of crops relies on the intricate knowledge of how the system works in the plant first. This thesis adds a small piece to the puzzle that is the plant immune system and allows a small step towards deciphering the plant immune response.

Appendix/Supplemental information

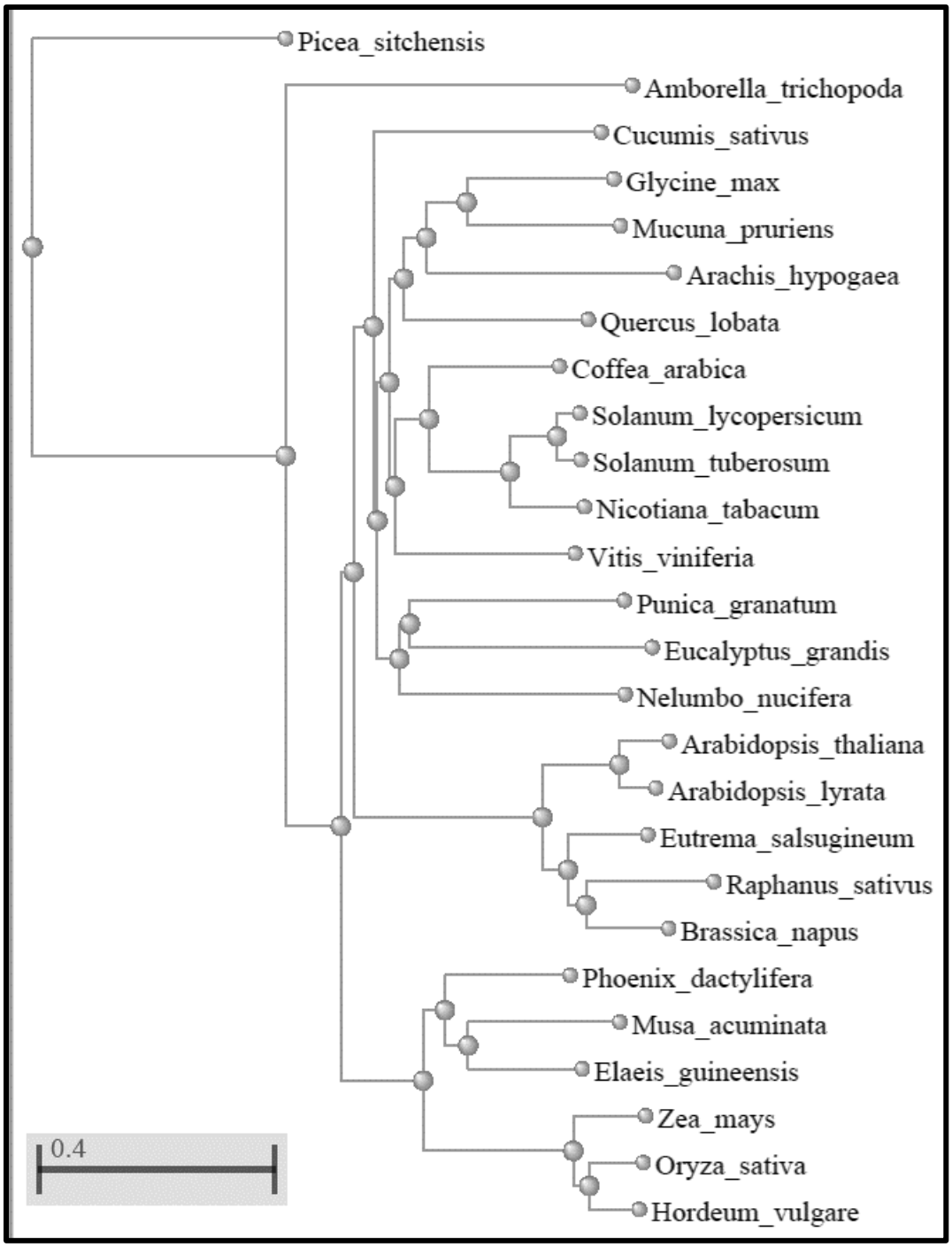

Figure S1: Phylogenetic tree constructed by analysis of the PAD4 orthologue sequences shown in figure 16. The tree was generated using the ClustalOmega tool.

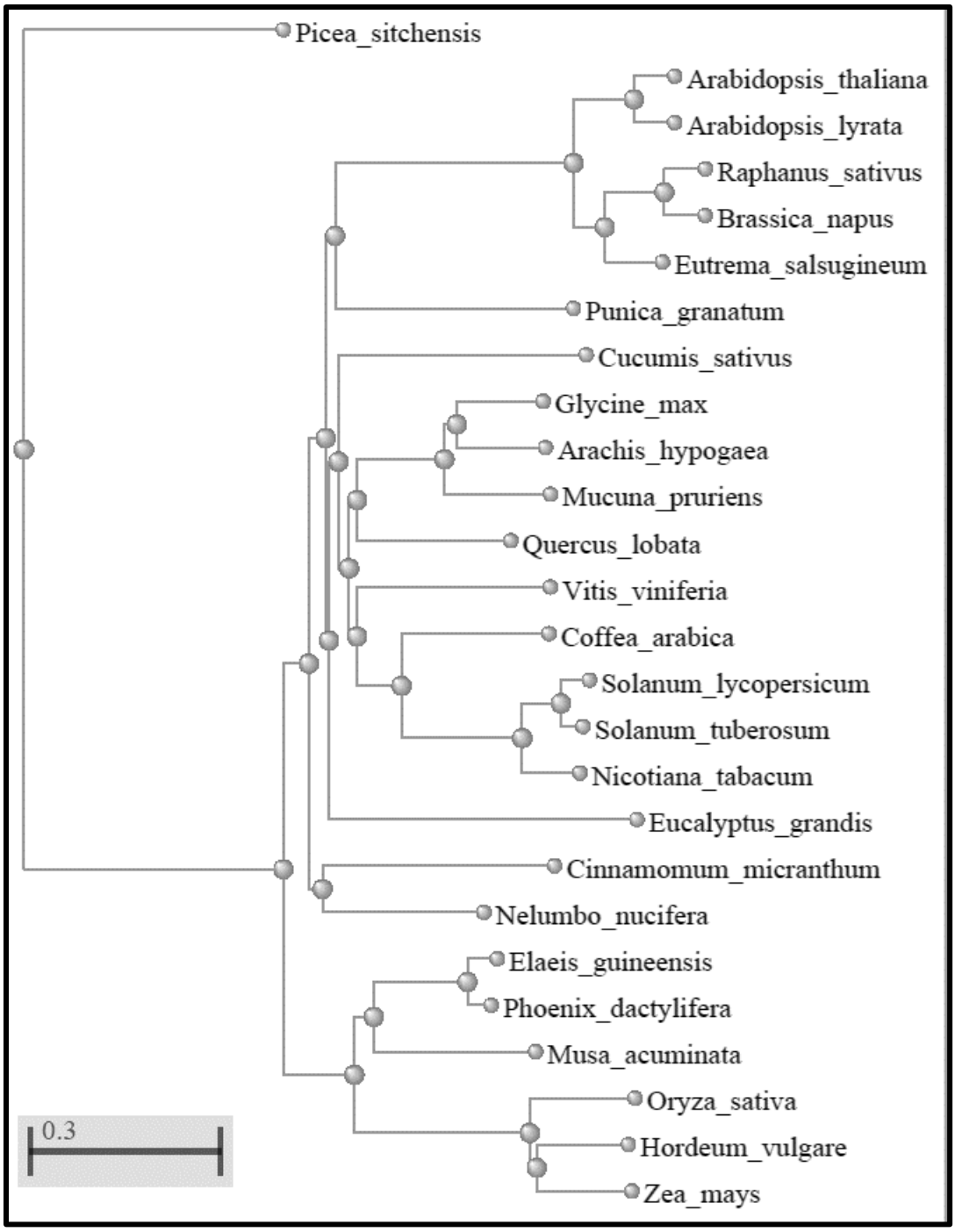

Figure S2 Phylogenetic tree constructed by analysis of the EDS1 orthologue sequences shown in figure 16. The tree was generated using the ClustalOmega tool.

References

Andexer, Jennifer; Langermann, Jan von; Mell, Annett; Bocola, Marco; Kragl, Udo; Eggert, Thorsten; Pohl, Martina (2007): An R-selective hydroxynitrile lyase from Arabidopsis thaliana with an alpha/beta-hydrolase fold. In: *Angewandte Chemie (International ed. in English)* 46 (45), S. 8679–8681. DOI: 10.1002/anie.200701455.

Auldridge, Michele E.; Guo, Yongxia; Austin, Michael B.; Ramsey, Justin; Fridman, Eyal; Pichersky, Eran; Noel, Joseph P. (2012): Emergent decarboxylase activity and attenuation of α/β-hydrolase activity during the evolution of methylketone biosynthesis in tomato. In: *The Plant cell* 24 (4), S. 1596–1607. DOI: 10.1105/tpc.111.093997.

Baggs, Erin L.; Monroe, J. Grey; Thanki, Anil S.; O'Grady, Ruby; Schudoma, Christian; Haerty, Wilfried; Krasileva, Ksenia V. (2020): Convergent Loss of an EDS1/PAD4 Signaling Pathway in Several Plant Lineages Reveals Coevolved Components of Plant Immunity and Drought Response. In: *The Plant cell* 32 (7), S. 2158–2177. DOI: 10.1105/tpc.19.00903.

Bernhofer, Michael; Dallago, Christian; Karl, Tim; Satagopam, Venkata; Heinzinger, Michael; Littmann, Maria et al. (2021): PredictProtein - Predicting Protein Structure and Function for 29 Years. In: *Nucleic acids research* 49 (W1), W535-W540. DOI: 10.1093/nar/gkab354.

Bhandari, Deepak D.; Lapin, Dmitry; Kracher, Barbara; Born, Patrick von; Bautor, Jaqueline; Niefind, Karsten; Parker, Jane E. (2019): An EDS1 heterodimer signalling surface enforces timely reprogramming of immunity genes in Arabidopsis. In: *Nature communications* 10 (1), S. 772. DOI: 10.1038/s41467-019-08783-0.

Bhattacharjee, Saikat; Halane, Morgan K.; Kim, Sang Hee; Gassmann, Walter (2011): Pathogen effectors target Arabidopsis EDS1 and alter its interactions with immune regulators. In: *Science (New York, N.Y.)* 334 (6061), S. 1405–1408. DOI: 10.1126/science.1211592.

Bi, Guozhi; Su, Min; Li, Nan; Liang, Yu; Dang, Song; Xu, Jiachao et al. (2021): The ZAR1 resistosome is a calcium-permeable channel triggering plant immune signaling. In: *Cell* 184 (13), 3528-3541.e12. DOI: 10.1016/j.cell.2021.05.003.

Bi, Guozhi; Su, Min; Li, Nan; Liang, Yu; Dang, Song; Xu, Jiachao et al. (2021): The ZAR1 resistosome is a calcium-permeable channel triggering plant immune signaling. In: *Cell* 184 (13), 3528-3541.e12. DOI: 10.1016/j.cell.2021.05.003.

Boivin, Stephane; Kozak, Sandra; Meijers, Rob (2013): Optimization of protein purification and characterization using Thermofluor screens. In: *Protein expression and purification* 91 (2), S. 192–206. DOI: 10.1016/j.pep.2013.08.002.

Brocchieri, Luciano; Karlin, Samuel (2005): Protein length in eukaryotic and prokaryotic proteomes. In: *Nucleic acids research* 33 (10), S. 3390–3400. DOI: 10.1093/nar/gki615.

Chang, Ming; Zhao, Jinping; Chen, Huan; Li, Guangyong; Chen, Jian; Li, Min et al. (2019): PBS3 Protects EDS1 from Proteasome-Mediated Degradation in Plant Immunity. In: *Molecular plant* 12 (5), S. 678–688. DOI: 10.1016/j.molp.2019.01.023.

Chen, Jian; Li, Min; Liu, Longyu; Chen, Gongyou; Fu, Zheng Qing (2021): ZAR1 resistosome and helper NLRs: Bringing in calcium and inducing cell death. In: *Molecular plant* 14 (8), S. 1234–1236. DOI: 10.1016/j.molp.2021.06.026.

Chong, Julie; Le Henanff, Gaëlle; Bertsch, Christophe; Walter, Bernard (2008): Identification, expression analysis and characterization of defense and signaling genes in Vitis vinifera. In: *Plant physiology and biochemistry : PPB* 46 (4), S. 469–481. DOI: 10.1016/j.plaphy.2007.09.010.

Dodds, Peter N.; Rathjen, John P. (2010): Plant immunity: towards an integrated view of plant-pathogen interactions. In: *Nature reviews. Genetics* 11 (8), S. 539–548. DOI: 10.1038/nrg2812.

Dongus, Joram A.; Bhandari, Deepak D.; Patel, Monika; Archer, Lani; Dijkgraaf, Lucas; Deslandes, Laurent et al. (2020): The Arabidopsis PAD4 Lipase-Like Domain Is Sufficient for Resistance to Green Peach Aphid. In: *Molecular plant-microbe interactions : MPMI* 33 (2), S. 328–335. DOI: 10.1094/MPMI-08-19-0245-R.

Dongus, Joram A.; Parker, Jane E. (2021): EDS1 signalling: At the nexus of intracellular and surface receptor immunity. In: *Current opinion in plant biology* 62, S. 102039. DOI: 10.1016/j.pbi.2021.102039.

Doye, Jonathan P. K.; Louis, Ard A.; Vendruscolo, Michele (2004): Inhibition of protein crystallization by evolutionary negative design. In: *Physical biology* 1 (1-2), P9-13. DOI: 10.1088/1478-3967/1/1/P02.

Duarte, Jose M.; Srebniak, Adam; Schärer, Martin A.; Capitani, Guido (2012): Protein interface classification by evolutionary analysis. In: *BMC bioinformatics* 13, S. 334. DOI: 10.1186/1471-2105-13-334.

Dupeux, Florine; Röwer, Martin; Seroul, Gael; Blot, Delphine; Márquez, José A. (2011): A thermal stability assay can help to estimate the crystallization likelihood of biological samples. In: *Acta crystallographica. Section D, Biological crystallography* 67 (Pt 11), S. 915–919. DOI: 10.1107/S0907444911036225.

Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. (2010): Features and development of Coot. In: *Acta crystallographica. Section D, Biological crystallography* 66 (Pt 4), S. 486–501. DOI: 10.1107/S0907444910007493.

Ericsson, Ulrika B.; Hallberg, B. Martin; Detitta, George T.; Dekker, Niek; Nordlund, Pär (2006): Thermofluor-based high-throughput stability optimization of proteins for structural studies. In: *Analytical biochemistry* 357 (2), S. 289–298. DOI: 10.1016/j.ab.2006.07.027.

Evans, Philip (2006): Scaling and assessment of data quality. In: *Acta Crystallogr D Biol Crystallogr* 62 (Pt 1), S. 72–82. DOI: 10.1107/S0907444905036693.

Evans, Philip R.; Murshudov, Garib N. (2013): How good are my data and what is the resolution? In: *Acta crystallographica. Section D, Biological crystallography* 69 (Pt 7), S. 1204–1214. DOI: 10.1107/S0907444913000061.

Falk, A.; Feys, B. J.; Frost, L. N.; Jones, J. D.; Daniels, M. J.; Parker, J. E. (1999): EDS1, an essential component of R gene-mediated disease resistance in Arabidopsis has homology to eukaryotic lipases. In: *Proceedings of the National Academy of Sciences of the United States of America* 96 (6), S. 3292–3297. DOI: 10.1073/pnas.96.6.3292.

Feys, B. J.; Moisan, L. J.; Newman, M. A.; Parker, J. E. (2001): Direct interaction between the Arabidopsis disease resistance signaling proteins, EDS1 and PAD4. In: *The EMBO journal* 20 (19), S. 5400–5411. DOI: 10.1093/emboj/20.19.5400.

Feys, Bart J.; Wiermer, Marcel; Bhat, Riyaz A.; Moisan, Lisa J.; Medina-Escobar, Nieves; Neu, Christina et al. (2005): Arabidopsis SENESCENCE-ASSOCIATED GENE101 stabilizes and signals within an ENHANCED DISEASE SUSCEPTIBILITY1 complex in plant innate immunity. In: *The Plant cell* 17 (9), S. 2601–2613. DOI: 10.1105/tpc.105.033910.

Gao, Fei; Dai, Ru; Pike, Sharon M.; Qiu, Wenping; Gassmann, Walter (2014): Functions of EDS1-like and PAD4 genes in grapevine defenses against powdery mildew. In: *Plant molecular biology* 86 (4-5), S. 381–393. DOI: 10.1007/s11103-014-0235-4.

Gao, Fei; Shu, Xiaomei; Ali, Mohammad Babar; Howard, Susanne; Li, Nan; Winterhagen, Patrick et al. (2010): A functional EDS1 ortholog is differentially regulated in powdery mildew resistant and susceptible grapevines and complements an Arabidopsis eds1 mutant. In: *Planta* 231 (5), S. 1037–1047. DOI: 10.1007/s00425-010-1107-z.

Gauhar, Aziz; Privezentzev, Cyril V.; Demydchuk, Mykhaylo; Gerlza, Tanja; Rieger, Julia; Kungl, Andreas J. et al. (2021): Single domain shark VNAR antibodies neutralize SARS-CoV-2 infection in vitro. In: *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 35 (11), e21970. DOI: 10.1096/fj.202100986RR.

Gloggnitzer, Jiradet; Akimcheva, Svetlana; Srinivasan, Arunkumar; Kusenda, Branislav; Riehs, Nina; Stampfl, Hansjörg et al. (2014): Nonsense-mediated mRNA decay modulates immune receptor levels to regulate plant antibacterial defense. In: *Cell host & microbe* 16 (3), S. 376–390. DOI: 10.1016/j.chom.2014.08.010.

Guo, Yongxia; Zheng, Zuyu; La Clair, James J.; Chory, Joanne; Noel, Joseph P. (2013): Smoke-derived karrikin perception by the α/β-hydrolase KAI2 from Arabidopsis. In: *Proceedings of the National Academy of Sciences of the United States of America* 110 (20), S. 8284–8289. DOI: 10.1073/pnas.1306265110.

HARTLEY, B. S.; KILBY, B. A. (1954): The reaction of p-nitrophenyl esters with chymotrypsin and insulin. In: *The Biochemical journal* 56 (2), S. 288–297. DOI: 10.1042/bj0560288.

Hassanzadeh-Ghassabeh, Gholamreza; Devoogdt, Nick; Pauw, Pieter de; Vincke, Cécile; Muyldermans, Serge (2013): Nanobodies and their potential applications. In: *Nanomedicine (London, England)* 8 (6), S. 1013–1026. DOI: 10.2217/nnm.13.86.

He, Yuehui; Gan, Susheng (2002): A gene encoding an acyl hydrolase is involved in leaf senescence in Arabidopsis. In: *The Plant cell* 14 (4), S. 805–815. DOI: 10.1105/tpc.010422.

Heidrich, Katharina; Wirthmueller, Lennart; Tasset, Céline; Pouzet, Cécile; Deslandes, Laurent; Parker, Jane E. (2011): Arabidopsis EDS1 connects pathogen effector recognition to cell compartment-specific immune responses. In: *Science (New York, N.Y.)* 334 (6061), S. 1401–1404. DOI: 10.1126/science.1211641.

Herzik, Mark A.; Wu, Mengyu; Lander, Gabriel C. (2019): High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM. In: *Nature communications* 10 (1), S. 1032. DOI: 10.1038/s41467-019-08991-8.

Hope, H. (1988): Cryocrystallography of biological macromolecules: a generally applicable method. In: *Acta crystallographica. Section B, Structural science* 44 (Pt 1), S. 22–26. DOI: 10.1107/s0108768187008632.

Horstman, Anneke; Tonaco, Isabella Antonia Nougalli; Boutilier, Kim; Immink, Richard G. H. (2014): A cautionary note on the use of split-YFP/BiFC in plant protein-protein interaction studies. In: *International journal of molecular sciences* 15 (6), S. 9628–9643. DOI: 10.3390/ijms15069628.

Huh, Sung Un; Cevik, Volkan; Ding, Pingtao; Duxbury, Zane; Ma, Yan; Tomlinson, Laurence et al. (2017): Protein-protein interactions in the RPS4/RRS1 immune receptor complex. In: *PLoS pathogens* 13 (5), e1006376. DOI: 10.1371/journal.ppat.1006376.

Jamil, Muhammad; Kountche, Boubacar A.; Al-Babili, Salim (2021): Current progress in Striga management. In: *Plant physiology* 185 (4), S. 1339–1352. DOI: 10.1093/plphys/kiab040.

Jeong, Hee-Jeong; Kim, Young Jin; Kim, Sang Hyon; Kim, Yoon-Ha; Lee, In-Jung; Kim, Yoon Ki; Shin, Jeong Sheop (2011): Nonsense-mediated mRNA decay factors, UPF1 and UPF3, contribute to plant defense. In: *Plant & cell physiology* 52 (12), S. 2147–2156. DOI: 10.1093/pcp/pcr144.

Jirage, D.; Tootle, T. L.; Reuber, T. L.; Frost, L. N.; Feys, B. J.; Parker, J. E. et al. (1999): Arabidopsis thaliana PAD4 encodes a lipase-like gene that is important for salicylic acid signaling. In: *Proceedings of the National Academy of Sciences of the United States of America* 96 (23), S. 13583–13588. DOI: 10.1073/pnas.96.23.13583.

Jones, Jonathan D. G.; Vance, Russell E.; Dangl, Jeffery L. (2016): Intracellular innate immune surveillance devices in plants and animals. In: *Science (New York, N.Y.)* 354 (6316). DOI: 10.1126/science.aaf6395.

Jubic, Lance M.; Saile, Svenja; Furzer, Oliver J.; El Kasmi, Farid; Dangl, Jeffery L. (2019): Help wanted: helper NLRs and plant immune responses. In: *Current opinion in plant biology* 50, S. 82–94. DOI: 10.1016/j.pbi.2019.03.013.

Jumper, John; Evans, Richard; Pritzel, Alexander; Green, Tim; Figurnov, Michael; Ronneberger, Olaf et al. (2021): Highly accurate protein structure prediction with AlphaFold. In: *Nature* 596 (7873), S. 583–589. DOI: 10.1038/s41586-021-03819-2.

Kabsch, Wolfgang (2010): XDS. In: *Acta crystallographica. Section D, Biological crystallography* 66 (Pt 2), S. 125–132. DOI: 10.1107/S0907444909047337.

Kalyna, Maria; Simpson, Craig G.; Syed, Naeem H.; Lewandowska, Dominika; Marquez, Yamile; Kusenda, Branislav et al. (2012): Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in Arabidopsis. In: *Nucleic acids research* 40 (6), S. 2454–2469. DOI: 10.1093/nar/gkr932.

Kelley, Lawrence A.; Mezulis, Stefans; Yates, Christopher M.; Wass, Mark N.; Sternberg, Michael J. E. (2015): The Phyre2 web portal for protein modeling, prediction and analysis. In: *Nature protocols* 10 (6), S. 845–858. DOI: 10.1038/nprot.2015.053.

Kim, Tae-Houn; Kunz, Hans-Henning; Bhattacharjee, Saikat; Hauser, Felix; Park, Jiyoung; Engineer, Cawas et al. (2012): Natural variation in small molecule-induced TIR-NB-LRR signaling induces root growth arrest via EDS1- and PAD4-complexed R protein VICTR in Arabidopsis. In: *The Plant cell* 24 (12), S. 5177–5192. DOI: 10.1105/tpc.112.107235.

Kirsch, Jack F.; Jencks, William P. (1964): Base Catalysis of Imidazole Catalysis of Ester Hydrolysis. In: *J. Am. Chem. Soc.* 86 (5), S. 833–837. DOI: 10.1021/ja01059a018.

Klimpel, Annika Präparation und Charakterisierung von „Phytoalexin Deficient 4“ (PAD4), eines Schlüsselproteins der pflanzlichen Immunantwort gegen pathogene Pilze und Insekten. Bachelorarbeit angefertig in der AG Niefind, Januar 2014.

Krissinel, Evgeny; Henrick, Kim (2007): Inference of macromolecular assemblies from crystalline state. In: *Journal of molecular biology* 372 (3), S. 774–797. DOI: 10.1016/j.jmb.2007.05.022.

Laemmli, U. K. (1970): Cleavage of structural proteins during the assembly of the head of bacteriophage T4. In: *Nature* 227 (5259), S. 680–685. DOI: 10.1038/227680a0.

Lapin, Dmitry; Bhandari, Deepak D.; Parker, Jane E. (2020): Origins and Immunity Networking Functions of EDS1 Family Proteins. In: *Annual review of phytopathology* 58, S. 253–276. DOI: 10.1146/annurev-phyto-010820-012840.

Lapin, Dmitry; Kovacova, Viera; Sun, Xinhua; Dongus, Joram A.; Bhandari, Deepak; Born, Patrick von et al. (2019): A Coevolved EDS1-SAG101-NRG1 Module Mediates Cell Death Signaling by TIR-Domain Immune Receptors. In: *The Plant cell* 31 (10), S. 2430–2455. DOI: 10.1105/tpc.19.00118.

Lenfant, Nicolas; Hotelier, Thierry; Velluet, Eric; Bourne, Yves; Marchot, Pascale; Chatonnet, Arnaud (2013): ESTHER, the database of the α/β-hydrolase fold superfamily of proteins: tools to explore diversity of functions. In: *Nucleic acids research* 41 (Database issue), D423-9. DOI: 10.1093/nar/gks1154.

Li, Qi; Wang, Ji; Bai, Tian; Zhang, Ming; Jia, Yuling; Shen, Danyu et al. (2020): A Phytophthora capsici effector suppresses plant immunity via interaction with EDS1. In: *Molecular plant pathology* 21 (4), S. 502–511. DOI: 10.1111/mpp.12912.

Li, Yuge; Yang, Yuhua; Hu, Yilong; Liu, Hailun; He, Ming; Yang, Ziyin et al. (2019): DELLA and EDS1 Form a Feedback Regulatory Module to Fine-Tune Plant Growth-Defense Tradeoff in Arabidopsis. In: *Molecular plant* 12 (11), S. 1485–1498. DOI: 10.1016/j.molp.2019.07.006.

Liebschner, Dorothee; Afonine, Pavel V.; Baker, Matthew L.; Bunkóczi, Gábor; Chen, Vincent B.; Croll, Tristan I. et al. (2019): Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. In: *Acta Crystallogr D Struct Biol* 75 (Pt 10), S. 861–877. DOI: 10.1107/S2059798319011471.

Liu, Hailun; Li, Yuge; Hu, Yilong; Yang, Yuhua; Zhang, Wenbin; He, Ming et al. (2021): EDS1-interacting J protein 1 is an essential negative regulator of plant innate immunity in Arabidopsis. In: *The Plant cell* 33 (1), S. 153–171. DOI: 10.1093/plcell/koaa007.

Louis, Joe; Gobbato, Enrico; Mondal, Hossain A.; Feys, Bart J.; Parker, Jane E.; Shah, Jyoti (2012): Discrimination of Arabidopsis PAD4 activities in defense against green peach aphid and pathogens. In: *Plant physiology* 158 (4), S. 1860–1872. DOI: 10.1104/pp.112.193417.

Louis, Joe; Shah, Jyoti (2015): Plant defence against aphids: the PAD4 signalling nexus. In: *Journal of experimental botany* 66 (2), S. 449–454. DOI: 10.1093/jxb/eru454.

Ma, Shoucai; Lapin, Dmitry; Liu, Li; Sun, Yue; Song, Wen; Zhang, Xiaoxiao et al. (2020): Direct pathogen-induced assembly of an NLR immune receptor complex to form a holoenzyme. In: *Science (New York, N.Y.)* 370 (6521). DOI: 10.1126/science.abe3069.

Marzec, Marek; Brewer, Philip (2019): Binding or Hydrolysis? How Does the Strigolactone Receptor Work? In: *Trends in plant science* 24 (7), S. 571–574. DOI: 10.1016/j.tplants.2019.05.001.

McCoy, Airlie J.; Grosse-Kunstleve, Ralf W.; Adams, Paul D.; Winn, Martyn D.; Storoni, Laurent C.; Read, Randy J. (2007): Phaser crystallographic software. In: *J Appl Crystallogr* 40 (Pt 4), S. 658–674. DOI: 10.1107/S0021889807021206.

McMahon C, Baier AS, Pascolutti R, Wegrecki M, Zheng S, Ong JX, Erlandson SC, Hilger D, Rasmussen SGF, Ring AM, Manglik A, Kruse AC. Yeast surface display platform for rapid discovery of conformationally selective nanobodies. Nat Struct Mol Biol. 2018 Mar;25(3):289-296. doi: 10.1038/s41594-018-0028-6

Meyers, Robert A. (Hg.) (2006): Encyclopedia of Analytical Chemistry. Chichester, UK: Wiley.

Mindrebo, Jeffrey T.; Nartey, Charisse M.; Seto, Yoshiya; Burkart, Michael D.; Noel, Joseph P. (2016): Unveiling the functional diversity of the alpha/beta hydrolase superfamily in the plant kingdom. In: *Current opinion in structural biology* 41, S. 233–246. DOI: 10.1016/j.sbi.2016.08.005.

Mistry, Jaina; Chuguransky, Sara; Williams, Lowri; Qureshi, Matloob; Salazar, Gustavo A.; Sonnhammer, Erik L. L. et al. (2021): Pfam: The protein families database in 2021. In: *Nucleic acids research* 49 (D1), D412-D419. DOI: 10.1093/nar/gkaa913.

Mullis, K.; Faloona, F.; Scharf, S.; Saiki, R.; Horn, G.; Erlich, H. (1986): Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. In: *Cold Spring Harbor symposia on quantitative biology* 51 Pt 1, S. 263–273. DOI: 10.1101/sqb.1986.051.01.032.

Murase, Kohji; Hirano, Yoshinori; Sun, Tai-ping; Hakoshima, Toshio (2008): Gibberellin-induced DELLA recognition by the gibberellin receptor GID1. In: *Nature* 456 (7221), S. 459–463. DOI: 10.1038/nature07519.

Muyldermans, Serge (2013): Nanobodies: natural single-domain antibodies. In: *Annual review of biochemistry* 82, S. 775–797. DOI: 10.1146/annurev-biochem-063011-092449.

Narusaka, Mari; Shirasu, Ken; Noutoshi, Yoshiteru; Kubo, Yasuyuki; Shiraishi, Tomonori; Iwabuchi, Masaki; Narusaka, Yoshihiro (2009): RRS1 and RPS4 provide a dual Resistance-gene system against fungal and bacterial pathogens. In: *The Plant journal : for cell and molecular biology* 60 (2), S. 218–226. DOI: 10.1111/j.1365-313X.2009.03949.x.

Nedrud, David M.; Lin, Hui; Lopez, Gilsinia; Padhi, Santosh K.; Legatt, Graig A.; Kaz-Lauskas, Romas J. (2014): Uncovering divergent evolution of α/β-hydrolases: a surprising residue substitution needed to convert Hevea brasiliensis hydroxynitrile lyase into an esterase. In: *Chemical science* 5 (11), S. 4265–4277. DOI: 10.1039/c4sc01544d.

Neubauer, Matthew; Serrano, Irene; Rodibaugh, Natalie; Bhandari, Deepak D.; Bautor, Jaqueline; Parker, Jane E.; Innes, Roger W. (2020): Arabidopsis EDR1 Protein Kinase Regulates the Association of EDS1 and PAD4 to Inhibit Cell Death. In: *Molecular plant-microbe interactions : MPMI* 33 (4), S. 693–703. DOI: 10.1094/MPMI-12-19-0339-R.

Ollis, D. L.; Cheah, E.; Cygler, M.; Dijkstra, B.; Frolow, F.; Franken, S. M. et al. (1992): The alpha/beta hydrolase fold. In: *Protein engineering* 5 (3), S. 197–211. DOI: 10.1093/protein/5.3.197.

Østdal, H.; Andersen, H. J. (1996): Non-enzymic protein induced hydrolysis of P-nitrophenyl acyl esters in relation to lipase/esterase assays. In: *Food Chemistry* 55 (1), S. 55–61. DOI: 10.1016/0308-8146(95)00094-1.

Padhi, Santosh Kumar; Fujii, Ryota; Legatt, Graig A.; Fossum, Sara L.; Berchtold, Reto; Kazlauskas, Romas J. (2010): Switching from an esterase to a hydroxynitrile lyase mechanism requires only two amino acid substitutions. In: *Chemistry & biology* 17 (8), S. 863–871. DOI: 10.1016/j.chembiol.2010.06.013.

Park, Yangshin; Shin, Inchul; Rhee, Sangkee (2015): Crystal structure of the effector protein HopA1 from Pseudomonas syringae. In: *Journal of structural biology* 189 (3), S. 276–280. DOI: 10.1016/j.jsb.2015.02.002.

PARK, J. H.; MERIWETHER, B. P.; CLODFELDER, P.; CUNNINGHAM, L. W. (1961): The hydrolysis of p-nitrophenyl acetate catalyzed by 3-phosphoglyceraldehyde dehydrogenase. In: *The Journal of biological chemistry* 236, S. 136–141.

Parker, J. E.; Holub, E. B.; Frost, L. N.; Falk, A.; Gunn, N. D.; Daniels, M. J. (1996): Characterization of eds1, a mutation in Arabidopsis suppressing resistance to Peronospora parasitica specified by several different RPP genes. In: *The Plant cell* 8 (11), S. 2033–2046. DOI: 10.1105/tpc.8.11.2033.

Rasmussen, Søren G. F.; DeVree, Brian T.; Zou, Yaozhong; Kruse, Andrew C.; Chung, Ka Young; Kobilka, Tong Sun et al. (2011): Crystal structure of the β2 adrenergic receptor-Gs protein complex. In: *Nature* 477 (7366), S. 549–555. DOI: 10.1038/nature10361.

Rauwerdink, Alissa; Kazlauskas, Romas J. (2015): How the Same Core Catalytic Machinery Catalyzes 17 Different Reactions: the Serine-Histidine-Aspartate Catalytic Triad of α/β-Hydrolase Fold Enzymes. In: *ACS catalysis* 5 (10), S. 6153–6176. DOI: 10.1021/acscatal.5b01539.

Rayson, Samantha; Arciga-Reyes, Luis; Wootton, Lucie; Torres Zabala, Marta de; Truman, William; Graham, Neil et al. (2012): A role for nonsense-mediated mRNA decay in plants: pathogen responses are induced in Arabidopsis thaliana NMD mutants. In: *PloS one* 7 (2), e31917. DOI: 10.1371/journal.pone.0031917.

Ricci, Daniela; Gidalevitz, Tali; Argon, Yair (2021): The special unfolded protein response in plasma cells. In: *Immunological reviews* 303 (1), S. 35–51. DOI: 10.1111/imr.13012.

Riehs-Kearnan, Nina; Gloggnitzer, Jiradet; Dekrout, Bettina; Jonak, Claudia; Riha, Karel (2012): Aberrant growth and lethality of Arabidopsis deficient in nonsense-mediated RNA decay factors is caused by autoimmune-like response. In: *Nucleic acids research* 40 (12), S. 5615–5624. DOI: 10.1093/nar/gks195.

Rosenbaum, Daniel M.; Zhang, Cheng; Lyons, Joseph A.; Holl, Ralph; Aragao, David; Arlow, Daniel H. et al. (2011): Structure and function of an irreversible agonist-β(2) adrenoceptor complex. In: *Nature* 469 (7329), S. 236–240. DOI: 10.1038/nature09665.

Schmidt, Andrea; Gruber, Karl; Kratky, Christoph; Lamzin, Victor S. (2008): Atomic resolution crystal structures and quantum chemistry meet to reveal subtleties of hydroxynitrile lyase catalysis. In: *The Journal of biological chemistry* 283 (31), S. 21827–21836. DOI: 10.1074/jbc.M801056200.

Schönherr, Robert; Rudolph, Janine Mia; Redecke, Lars (2018): Protein crystallization in living cells. In: *Biological chemistry* 399 (7), S. 751–772. DOI: 10.1515/hsz-2018-0158.

Schoof, Michael; Faust, Bryan; Saunders, Reuben A.; Sangwan, Smriti; Rezelj, Veronica; Hoppe, Nick et al. (2020): An ultrapotent synthetic nanobody neutralizes SARS-CoV-2 by stabilizing inactive Spike. In: *Science (New York, N.Y.)* 370 (6523), S. 1473–1479. DOI: 10.1126/science.abe3255.

Seto, Yoshiya; Yasui, Rei; Kameoka, Hiromu; Tamiru, Muluneh; Cao, Mengmeng; Terauchi, Ryohei et al. (2019): Strigolactone perception and deactivation by a hydrolase receptor DWARF14. In: *Nature communications* 10 (1), S. 191. DOI: 10.1038/s41467-018-08124-7.

Shimada, Asako; Ueguchi-Tanaka, Miyako; Nakatsu, Toru; Nakajima, Masatoshi; Naoe, Youichi; Ohmiya, Hiroko et al. (2008): Structural basis for gibberellin recognition by its receptor GID1. In: *Nature* 456 (7221), S. 520–523. DOI: 10.1038/nature07546.

Sievers, Fabian; Wilm, Andreas; Dineen, David; Gibson, Toby J.; Karplus, Kevin; Li, Weizhong et al. (2011): Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. In: *Molecular systems biology* 7, S. 539. DOI: 10.1038/msb.2011.75.

Slabinski, Lukasz; Jaroszewski, Lukasz; Rychlewski, Leszek; Wilson, Ian A.; Lesley, Scott A.; Godzik, Adam (2007): XtalPred: a web server for prediction of protein crystallizability. In: *Bioinformatics (Oxford, England)* 23 (24), S. 3403–3405. DOI: 10.1093/bioinformatics/btm477.

Spoel, Steven H.; Dong, Xinnian (2012): How do plants achieve immunity? Defence without specialized immune cells. In: *Nature reviews. Immunology* 12 (2), S. 89–100. DOI: 10.1038/nri3141.

Sun, Xinhua; Lapin, Dmitry; Feehan, Joanna M.; Stolze, Sara C.; Kramer, Katharina; Dongus, Joram A. et al. (2021): Pathogen effector recognition-dependent association of NRG1 with EDS1 and SAG101 in TNL receptor immunity. In: *Nature communications* 12 (1), S. 3335. DOI: 10.1038/s41467-021-23614-x.

Terwilliger, Thomas C.; Grosse-Kunstleve, Ralf W.; Afonine, Pavel V.; Moriarty, Nigel W.; Zwart, Peter H.; Hung, Li Wei et al. (2008): Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. In: *Acta Crystallogr D Biol Crystallogr* 64 (Pt 1), S. 61–69. DOI: 10.1107/S090744490705024X.

Traenkle, Bjoern; Kaiser, Philipp D.; Rothbauer, Ulrich (2006): Nanobody Platform for Determination of Protein Structure, Application of. In: Robert A. Meyers (Hg.): Encyclopedia of Analytical Chemistry. Chichester, UK: Wiley, S. 1–14.

Vander Meersche, Yann; Cretin, Gabriel; Brevern, Alexandre G. de; Gelly, Jean-Christophe; Galochkina, Tatiana (2021): MEDUSA: Prediction of Protein Flexibility from Sequence. In: *Journal of molecular biology* 433 (11), S. 166882. DOI: 10.1016/j.jmb.2021.166882.

Vedadi, Masoud; Niesen, Frank H.; Allali-Hassani, Abdellah; Fedorov, Oleg Y.; Finerty, Patrick J.; Wasney, Gregory A. et al. (2006): Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. In: *Proceedings of the National Academy of Sciences of the United States of America* 103 (43), S. 15835–15840. DOI: 10.1073/pnas.0605224103.

Vonrhein, Clemens; Flensburg, Claus; Keller, Peter; Sharff, Andrew; Smart, Oliver; Paciorek, Wlodek et al. (2011): Data processing and analysis with the autoPROC toolbox. In: *Acta crystallographica. Section D, Biological crystallography* 67 (Pt 4), S. 293–302. DOI: 10.1107/S0907444911007773.

Voss, Martin; Toelzer, Christine; Bhandari, Deepak D.; Parker, Jane E.; Niefind, Karsten (2019): Arabidopsis immunity regulator EDS1 in a PAD4/SAG101-unbound form is a monomer with an inherently inactive conformation. In: *Journal of structural biology* 208 (3), S. 107390. DOI: 10.1016/j.jsb.2019.09.007.

Wagner, Stephan; Stuttmann, Johannes; Rietz, Steffen; Guerois, Raphael; Brunstein, Elena; Bautor, Jaqueline et al. (2013): Structural basis for signaling by exclusive EDS1 heteromeric complexes with SAG101 or PAD4 in plant innate immunity. In: *Cell host & microbe* 14 (6), S. 619–630. DOI: 10.1016/j.chom.2013.11.006.

Wagner, Stephan (2013). *Analysing Novel Structures of Protein Complexes Requiring the Central Plant Immune Regulator EDS1.* Dissertation, Universität zu Köln

Wan, Li; Essuman, Kow; Anderson, Ryan G.; Sasaki, Yo; Monteiro, Freddy; Chung, Eui-Hwan et al. (2019): TIR domains of plant immune receptors are NAD+-cleaving enzymes that promote cell death. In: *Science (New York, N.Y.)* 365 (6455), S. 799–803. DOI: 10.1126/science.aax1771.

Wan, Li; Essuman, Kow; Anderson, Ryan G.; Sasaki, Yo; Monteiro, Freddy; Chung, Eui-Hwan et al. (2019): TIR domains of plant immune receptors are NAD+-cleaving enzymes that promote cell death. In: *Science (New York, N.Y.)* 365 (6455), S. 799–803. DOI: 10.1126/science.aax1771.

Wang, Jialin; Shine, M. B.; Gao, Qing-Ming; Navarre, Duroy; Jiang, Wei; Liu, Chunyan et al. (2014): Enhanced Disease Susceptibility1 Mediates Pathogen Resistance and Virulence Function of a Bacterial Effector in Soybean. In: *Plant physiology* 165 (3), S. 1269–1284. DOI: 10.1104/pp.114.242495.

Wang, Jizong; Hu, Meijuan; Wang, Jia; Qi, Jinfeng; Han, Zhifu; Wang, Guoxun et al. (2019): Reconstitution and structure of a plant NLR resistosome conferring immunity. In: *Science (New York, N.Y.)* 364 (6435). DOI: 10.1126/science.aav5870.

Wäspi, U.; Misteli, B.; Hasslacher, M.; Jandrositz, A.; Kohlwein, S. D.; Schwab, H.; Dudler, R. (1998): The defense-related rice gene Pir7b encodes an alpha/beta hydrolase fold protein exhibiting esterase activity towards naphthol AS-esters. In: *European journal of biochemistry* 254 (1), S. 32–37. DOI: 10.1046/j.1432-1327.1998.2540032.x.

Westfield, Gerwin H.; Rasmussen, Søren G. F.; Su, Min; Dutta, Somnath; DeVree, Brian T.; Chung, Ka Young et al. (2011): Structural flexibility of the G alpha s alpha-helical domain in the beta2-adrenoceptor Gs complex. In: *Proceedings of the National Academy of Sciences of the United States of America* 108 (38), S. 16086–16091. DOI: 10.1073/pnas.1113645108.

Wiermer, Marcel; Feys, Bart J.; Parker, Jane E. (2005): Plant immunity: the EDS1 regulatory node. In: *Current opinion in plant biology* 8 (4), S. 383–389. DOI: 10.1016/j.pbi.2005.05.010.

Wilkins, M. R.; Gasteiger, E.; Bairoch, A.; Sanchez, J. C.; Williams, K. L.; Appel, R. D.; Hochstrasser, D. F. (1999): Protein identification and analysis tools in the ExPASy server. In: *Methods in molecular biology (Clifton, N.J.)* 112, S. 531–552. DOI: 10.1385/1-59259-584-7:531.

Zavrtanik, Uroš; Lukan, Junoš; Loris, Remy; Lah, Jurij; Hadži, San (2018): Structural Basis of Epitope Recognition by Heavy-Chain Camelid Antibodies. In: *Journal of molecular biology* 430 (21), S. 4369–4386. DOI: 10.1016/j.jmb.2018.09.002.

Zhang, Yunwei; Bai, Yang; Wu, Guangheng; Zou, Shenghao; Chen, Yongfang; Gao, Caixia; Tang, Dingzhong (2017): Simultaneous modification of three homoeologs of TaEDR1 by genome editing enhances powdery mildew resistance in wheat. In: *The Plant journal : for cell and molecular biology* 91 (4), S. 714–724. DOI: 10.1111/tpj.13599.

Zhu, Haocheng; Li, Chao; Gao, Caixia (2020): Applications of CRISPR-Cas in agriculture and plant biotechnology. In: *Nature reviews. Molecular cell biology* 21 (11), S. 661–677. DOI: 10.1038/s41580-020-00288-9.

„Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

Hürth, den 03.02.2022

Martin Voß